SIAM J. Math. Anal.
Vol. 54, No. 2, pp. 2251–2294

A MCKEAN–VLASOV SDE AND PARTICLE SYSTEM WITH INTERACTION FROM REFLECTING BOUNDARIES*

MICHELE COGHI†, WOLFGANG DREYER‡, PETER K. FRIZ§, PAUL GAJEWSKI‡, CLEMENS GUHLKE‡, AND MARIO MAURELLI¶

Abstract. We consider a one-dimensional McKean–Vlasov SDE on a domain and the associated mean-field interacting particle system. The peculiarity of this system is the combination of the interaction, which keeps the average position prescribed, and the reflection at the boundaries; these two factors make the effect of reflection nonlocal. We show pathwise well-posedness for the McKean–Vlasov SDE and convergence for the particle system in the limit of large particle number.

Key words. McKean–Vlasov SDEs, mean-field interacting diffusions, reflecting boundaries

AMS subject classifications. 60H10, 60K35, 60F99

DOI. 10.1137/21M1409421

1. Introduction.

In this paper, we consider a system of N interacting one-dimensional diffusions, with the following two main features: (a) they are confined in a bounded domain with reflecting boundaries; (b) their empirical average is prescribed. We show that, as the number of particles N goes to infinity, the system converges to the unique solution to a suitable McKean–Vlasov SDE on the domain.

The model and the results. The propotypical example is the following one:

$$\begin{aligned} &\mathrm{d}X_t^{i,N} = \mathrm{d}W_t^i + \mathrm{d}K_t^N - \mathrm{d}k_t^{i,N}, \\ &X_t^{i,N} \in [0,1], \quad \mathrm{d}k_t^{i,N} = n(X_t^{i,N})\mathrm{d}|k^{i,N}|_t, \quad \mathrm{d}|k^{i,N}|_t = 1_{X_t^{i,N}\in\{0,1\}}\mathrm{d}|k^{i,N}|_t, \\ &\frac{1}{N}\sum_{i=1}^N X_t^{i,N} = q. \end{aligned} \tag{1.1}$$

Here W^i are independent real Brownian motions, q is the given average in $(0,1)$, n is the outer normal on $\partial[0,1] = \{0,1\}$, and the solution is a triple $X^{(N)} = (X^{i,N})_{i=1,\dots N}$, $k^{(N)} = (k^{i,N})_{i=1,\dots N}$, K^N satisfying the above system; $|k^{i,N}|$ denotes the total variation process of $k^{i,N}$. We will sometimes omit the superscript N from the notation.

*Received by the editors April 8, 2021; accepted for publication (in revised form) November 16, 2021; published electronically April 19, 2022.
https://doi.org/10.1137/21M1409421

Funding: The work of the authors was supported by the Einstein Center for Mathematics Berlin through MATHEON projects C-SE8 and C-SE17, "Stochastic methods for the analysis of lithium-ion batteries." The work of the first and sixth authors was supported by the Hausdorff Research Institute for Mathematics in Bonn under the Junior Trimester Program "Randomness, PDEs and Nonlinear Fluctuations." The work of the third author was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant 683164.

†Dipartimento di Matematica, Università degli Studi di Trento, via Sommarive 14, 38123 Povo (Trento), Italy (michele.coghi@unitn.it).

‡Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstraße 39, 10117 Berlin, Germany (Wolfgang.Dreyer@wias-berlin.de, gajewskp@wias-berlin.de, guhlke@wias-berlin.de).

§Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstraße 39, 10117 Berlin, Germany, and Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany (p.k.friz@gmail.com).

¶Dipartimento di Matematica, Università degli Studi di Milano, via Saldini 50, 20133 Milano, Italy (mario.maurelli@unimi.it).

The term $-k^i$ represents the reflection of the process X^i at the boundary of $[0,1]$ and the term K (independent of i) represents the interaction between the particles, which keeps the average equal to q. More generally, for modeling purposes, we consider also the case with a given drift $\mu : [0,1] \to \mathbb{R}$, a given time-dependent average $q : [0,T] \to (0,1)$, and a constant noise intensity $\sigma \in \mathbb{R}$, namely we take the system

$$\begin{aligned} &\mathrm{d}X_t^{i,N} = -\mu(X_t^{i,N})\mathrm{d}t + \sigma \mathrm{d}W_t^i + \mathrm{d}K_t^N - \mathrm{d}k_t^{i,N}, \\ &X_t^{i,N} \in [0,1], \quad \mathrm{d}k_t^{i,N} = n(X_t^{i,N})\mathrm{d}|k^{i,N}|_t, \quad \mathrm{d}|k^{i,N}|_t = 1_{X_t^{i,N} \in \{0,1\}}\mathrm{d}|k^{i,N}|_t, \\ &\frac{1}{N}\sum_{i=1}^N X_t^{i,N} = q(t). \end{aligned} \tag{1.2}$$

The last line of the above system can be easily converted into an expression for K^N in terms of $X^{(N)}$ and $k^{(N)}$, namely

$$\mathrm{d}K_t^N = \left(\frac{1}{N}\sum_{i=1}^N \mu(X_t^i) + \dot{q}(t)\right)\mathrm{d}t - \sigma\frac{1}{N}\sum_{i=1}^N \mathrm{d}W_t^i + \frac{1}{N}\sum_{i=1}^N \mathrm{d}k_t^i. \tag{1.3}$$

The main novelty of this work is the peculiar combination of the reflecting boundary and the condition on the average of the particles. This combination is reflected in formula (1.3), where the interaction $\mathrm{d}K$ depends also on the empirical average of $\mathrm{d}k^i$. To guess the limiting behavior (as $N \to \infty$) of the system (1.2), we can replace the average over particles $N^{-1}\sum_{i=1}^N$ with the average over the probability space $\mathbb{E}$. In this way, we get the following McKean–Vlasov SDE on the domain $[0,1]$:

$$\begin{aligned} &\mathrm{d}\bar{X}_t = -\mu(\bar{X}_t)\mathrm{d}t + \sigma\mathrm{d}W_t + \mathrm{d}\bar{K}_t - \mathrm{d}\bar{k}_t, \\ &\bar{X}_t \in [0,1], \quad \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t, \quad \mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t \in \{0,1\}}\mathrm{d}|\bar{k}|_t, \\ &\mathbb{E}\bar{X}_t = q(t), \end{aligned} \tag{1.4}$$

where W is a real Brownian motion and the solution is a triple $\bar{X}, \bar{k}, \bar{K}$ satisfying the above SDE. As in the particle system, the last line of (1.4) can be converted into an expression for $\bar{K}$:

$$\mathrm{d}\bar{K}_t = (\mathbb{E}\mu(\bar{X}_t) + \dot{q}(t))\mathrm{d}t + \mathbb{E}\mathrm{d}\bar{k}_t. \tag{1.5}$$

Our main results, Theorems 3.9 and 3.10 and Proposition 3.12, state roughly speaking that the McKean–Vlasov SDE (1.4) is well-posed in the pathwise sense and the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ from the system (1.2) converges in probability, as $N \to \infty$, to the law of the unique solution to (1.4), with its time marginals converging in L^1.

Motivation. Our motivation to study this system comes from a specific model for charging and discharging in a lithium-ion battery, introduced in [DGH11] and further studied and expanded, for example, in [DHM$^+$15, GGM$^+$18]. In this model, roughly speaking, the lithium atoms enter and exit iron phosphate particles in the cathode. The Y_t^i represents the filling degree of the ith iron phosphate particle at time t (for example, $Y^i = 1$, resp., $= 0$, stands for the ith particle fully filled with lithium atoms, resp. fully empty,); by this definition of Y^i, Y^i has to stay in $[0,1]$. The prescribed average $q(t)$ of Y_t^i represents the current in the battery, which is given and is proportional to the percentage of lithium atoms inside the ensemble of particles. The reason to consider reflecting boundaries $Y^i \in \{0,1\}$ comes from the boundary conditions in the Fokker–Planck equation in [DHM+15]; this choice of boundary conditions is convenient mathematically, though the physical motivation is less clear. From a mathematical perspective, [DHM+15] shows global well-posedness for the nonlinear nonlocal Fokker–Planck equation associated with the McKean–Vlasov SDE (1.4), namely

$$
\begin{aligned}
&\partial_t u(t,x) + \partial_x[(-\mu(x) + \dot{\bar{K}}_t)u(t,x)] = \frac{\sigma^2}{2}\partial_x^2 u(t,x), \quad t>0,\, x\in(0,1),\\
&\frac{\sigma^2}{2}\partial_x u(t,x) + (\mu(x) - \dot{\bar{K}}_t)u(t,x) = 0, \quad t>0,\, x\in\partial(0,1), \qquad (1.6)\\
&\int_0^1 x u(t,x)\mathrm{d}x = q(t), \quad t\ge 0.
\end{aligned}
$$

The paper [GGM+18] considers the particle system (1.6), associated with (1.2), even in a more general version (to take into account variations in the radius of iron phosphate particles), but removes the boundaries: without boundaries, the particle system (1.2) is reduced to a classical system of mean-field interacting diffusions, for which convergence to the corresponding McKean–Vlasov SDE is well-known. Hence the current paper arises from the natural (from a mathematical viewpoint) question of whether convergence of the particle system for the model (1.6) holds. We also point out that interacting diffusions with constraints both on the domain and on the empirical measure of the diffusions appear in several contexts; see, e.g., [BCdRGL20, Jab17, Bar20] below.

Background. McKean–Vlasov SDEs are SDEs where the drift depends also on the law of the solution, namely SDEs of the form

$$
\mathrm{d}\bar{X}_t = b\left(\bar{X}_t, \mathrm{Law}(\bar{X}_t)\right)\mathrm{d}t + \mathrm{d}W_t, \qquad (1.7)
$$

where W is a given Brownian motion (we do not consider here the case of general diffusion coefficients). McKean–Vlasov SDEs are related to the mean-field interacting diffusions, namely systems of the form

$$
\mathrm{d}X_t^i = b\left(X_t^i, \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}\right)\mathrm{d}t + \mathrm{d}W_t^i, \quad i=1,\dots N,
$$

where W^i are independent Brownian motions. By classical resuls, e.g., [Szn91, M$\acute{9}$6, Tan84], if b is bounded and smooth (smoothness with respect to the measure variable is understood in the sense of Wasserstein distance), then the McKean–Vlasov SDE (1.7) is pathwise well-posed and, as $N\to\infty$, the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X^i}$ converges to the law of the solution $\bar{X}$ to the McKean–Vlasov SDE. This convergence result is a law of large numbers type result and is related to the asymptotic independence of the particles, the so-called propagation of chaos; see, e.g., [Szn84]. The Fokker–Planck equation associated with the McKean–Vlasov SDE (1.7), namely the equation for $\mathrm{Law}(\bar{X}_t)$, is nonlinear; see section 2.3.

SDEs on a domain $\bar{D}\subseteq\mathbb{R}^m$ with reflecting boundaries take the form

$$
\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t - \mathrm{d}k_t, \qquad (1.8)
$$

$$
X_t\in\bar{D}, \quad \mathrm{d}k_t = n(X_t)\mathrm{d}|k|_t, \quad \mathrm{d}|k|_t = 1_{X_t\in\partial D}\mathrm{d}|k|_t, \qquad (1.9)
$$

where W is a Brownian motion (we do not consider general diffusion coefficients) and $n(x)$ is the outer normal to D at x; $|k|$ represents the total variation process associated with k. The solution is a couple (X, k) and $-\mathrm{d}k$ represents a "kick," in the inward normal direction $-n(X)$, that the diffusion X receives anytime it reaches the boundary, and that makes X stay in the domain $\bar{D}$. Pathwise well-posedness for the SDE (1.8) has been proved under quite general conditions; see, e.g., [LS84, Tan79]. The Fokker–Planck equation associated with the SDE (1.8) has Neumann-type boundary conditions; see section 2.2.

To our knowledge, the first work to deal with both McKean–Vlasov SDEs and reflecting boundaries is [Szn84]: there pathwise well-posedness is proved for the SDE

$$\begin{aligned}
&\mathrm{d}\bar{X}_t = b(\bar{X}_t, \mathrm{Law}(\bar{X}_t))\mathrm{d}t + \mathrm{d}W_t - \mathrm{d}\bar{k}_t,\\
&\bar{X}_t \in \bar{D}, \quad \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t, \quad \mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t \in \partial D}\mathrm{d}|\bar{k}|_t,
\end{aligned}$$

and convergence is shown for the particle system

$$\begin{aligned}
&\mathrm{d}X^i_t = b\left(X^i_t, \frac{1}{N}\sum_{i=1}^N \delta_{X^i_t}\right)\mathrm{d}t + \mathrm{d}W^i_t - dk^i_t, \quad i = 1, \dots N,\\
&X^i_t \in \bar{D}, \quad \mathrm{d}k^i_t = n(X^i_t)\mathrm{d}|k^i|_t, \quad \mathrm{d}|k^i|_t = 1_{X^i_t \in \partial D}\mathrm{d}|k^i|_t, \quad i = 1, \dots N.
\end{aligned}$$

Other works have studied McKean–Vlasov SDEs with reflecting boundaries, in more general contexts, especially in the context of backward SDEs; see, e.g., [Li14]. However, in [Szn84] and in many of these works, the reflection is local, that is, at the level of the McKean–Vlasov SDE, $\mathrm{Law}(\mathrm{d}\bar{k}_t)$ does not appear in the SDE; at the level of the particle system, $\mathrm{d}k^i$ acts only on the ith particle X^i.

Closer to our work are the mean reflected (possibly backward) SDEs and related particle systems, introduced in [BEH18, BCdRGL20], and their generalization, namely the SDEs with a constraint on the law and related particle systems, introduced in [BCCdRH20]. Roughly speaking, in these SDEs a reflecting boundary is imposed on the law of the process. The typical example of this type of SDE is the following:

$$\begin{aligned}
&\mathrm{d}\bar{X}_t = b(\bar{X}_t, \mathrm{Law}(\bar{X}_t))\mathrm{d}t + \mathrm{d}W_t - \mathrm{d}\bar{K}_t,\\
&\mathrm{Law}(\bar{X}_t) \in \bar{\mathcal{D}},\ \mathrm{d}\bar{K}_t \text{ deterministic, "nonzero only when } \bar{X}_t \text{ is on the boundary of } \bar{\mathcal{D}}\text{."}
\end{aligned} \tag{1.10}$$

For such systems, under suitable conditions, [BCCdRH20] proves well-posedness and particle approximation. As a particular case, taking $\bar{\mathcal{D}} = \{\rho \mid \int x\rho(\mathrm{d}x) \geq q\}$ for a given $q \in \mathbb{R}$, the constraint becomes $\mathbb{E}[\bar{X}_t] \geq q$, which is essentially comparable to our constraint $\mathbb{E}[\bar{X}_t] = q(t)$ in the last line of (1.4). Due to the assumptions on $\bar{\mathcal{D}}$ (which must have a nonempty "interior"), the condition $\mathbb{E}[\bar{X}_t] = q(t)$ is not covered by [BCCdRH20], but this is not a big limitation: the condition $\mathbb{E}[\bar{X}_t] \geq q$ is actually more difficult to take into account than $\mathbb{E}[\bar{X}_t] = q(t)$, which gives an explicit form for $\bar{K}$ and makes the SDE a classical McKean–Vlasov SDE. However, compared to our equation (1.4), the restriction to deterministic $\bar{K}_t$ in [BCCdRH20] does not allow one to consider reflecting boundaries for the process $\bar{X}$ (for reflecting boundaries on $\bar{X}$, the reflection $\mathrm{d}\bar{k}$ is not deterministic). When one adds reflecting boundaries in [BCCdRH20], additional difficulties come into play; see Remark 4.1.

Probably the closest work to ours is [Jab17]. This paper considers a more general case than [BCCdRH20], in particular removing from (1.10) the requirement that $\bar{K}$ is deterministic. In particular, taking

$$\bar{\mathcal{D}} = \left\{\rho \mid \int x\rho(\mathrm{d}x) \geq q,\ \mathrm{supp}(\rho) \subseteq [0,1]\right\} \tag{1.11}$$

the constraint on $\bar{X}$ in (1.10) becomes

$$\mathbb{E}\bar{X}_t \geq q, \quad \bar{X}_t \in [0,1],$$

which is essentially similar to our constraints $\mathbb{E}\bar{X}_t = q(t)$, $\bar{X}_t \in [0,1]$ in (1.4). The work [Jab17] constructs a weak solution to the SDE (1.10) (without the requirement of deterministic $\bar{K}$, in particular including condition (1.11)) by a penalization approach. However, it does not show uniqueness, nor does it consider the related particle system.

The work [Bar20] studies a system of N Brownian particles X^i hitting a Newtonian moving barrier Y. For this system the paper proves the convergence, as $N \to \infty$, to a McKean–Vlasov type SDE, whose associated Fokker–Planck equation solves a free-boundary problem. Now, in the frame of the moving barrier, that is taking $Z^i_t = X^i_t - Y_t$, the system in [Bar20] is similar to our model (1.2), without the drift μ, but with one important difference: in the expression (1.3) for $\mathrm{d}K^N$, the term $\frac{1}{N}\sum_{i=1}^N \mathrm{d}k^i$ is replaced in [Bar20] by $\frac{1}{N}\sum_{i=1}^N k^i \mathrm{d}t$ (times some constant). In particular, unlike here, the term $\mathrm{d}K^N$ becomes of bounded variation in time in [Bar20].

The paper [DEH19] studies the case of backward SDEs with reflecting boundaries depending both on the diffusion process $\bar{X}$ and on the law of $\bar{X}$, showing well-posedness for this type of SDE and convergence for the corresponding penalization scheme. However, by the precise assumptions in [DEH19], a condition of the form $\mathbb{E}[\bar{X}_t] \geq q$ or $= q$ cannot be taken in [DEH19].

Finally, we mention the works [CCP11, DIRT15] and [HLSj19]: they deal with systems of interacting diffusions, which arise respectively in neuroscience and in finance, and include also a nonlocal effect of boundaries, though the boundaries are not reflecting. More precisely, when one or more particles hit the boundary, the other particles make a jump proportional to the number of particles hitting the boundary.

Novelty of our work. The main feature of our model is the combination of reflecting boundaries and nonlocal interaction. At the level of the McKean–Vlasov SDE (1.4), this combination appears in the formula (1.5) for the interacting term $\mathrm{d}\bar{K}$, which contains the term $\mathrm{d}\bar{k}$. At the level of the particle system (1.2), this fact corresponds to an oblique reflection for $X^{(N)} = (X^{1,N}, \dots X^{N,N})$, where the direction of reflection depends on the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_t}$. As explained before, to our knowledge, this kind of systems is studied only in [Jab17] (which shows an existence result for the McKean–Vlasov SDE).

More specifically, in our model (1.2), the nonlocal interaction comes from the condition $\frac{1}{N}\sum_{i=1}^N X^{i,N}_t = q(t)$ and keeps the direction of reflection for $X^{(N)}$ on the iperplane $\{x \mid \frac{1}{N}\sum_{i=1}^N x^i = 0\}$. Intuitively, when a particle X^i hits the boundary and receives a "kick" $-\mathrm{d}k^i$, then the other particles receive a kick $\frac{1}{N}\mathrm{d}k^j$ in the opposite direction, so that the average of the particles remains $q(t)$.

As we will explain below, our proof relies strongly on the constraint $\frac{1}{N}\sum_{i=1}^N X^{i,N}_t = q(t)$ and so on the specific form of the interaction. The question of well-posedness and particle approximation for a more general dependence of $\mathrm{d}\bar{K}$ on $\mathrm{d}\bar{k}$, or equivalently, of the direction of reflection of $X^{(N)}$ on the empirical measure, remains open.

Method of proof. The main idea of the proof is that both $\mathrm{d}\bar{k}$ and $\mathrm{d}\bar{K}$ in (1.4) act as projectors. Precisely, $\mathrm{d}\bar{k}$ acts as projector on the set of paths staying in $[0,1]$: indeed, if $(\bar{X}, \bar{k}^{\bar{X}}, \bar{K}^{\bar{X}})$ and $(\bar{X}, \bar{k}^{\bar{X}}, \bar{K}^{\bar{X}})$ are two solutions to (1.4), then $(\bar{X} - \bar{Y}) \cdot \mathrm{d}\bar{k}^{\bar{X}} \leq 0$. The term $\mathrm{d}\bar{K}$ acts as projector in $L^2(\Omega)$ (where $(\Omega, \mathcal{A}, P)$ is the underlying probability space) on the space of processes Z with average $\mathbb{E}[Z_t] = q(t)$: indeed, if $(\bar{X}, \bar{k}^{\bar{X}}, \bar{K}^{\bar{X}})$, $(\bar{X}, \bar{k}^{\bar{X}}, \bar{K}^{\bar{X}})$ are two solutions, then $\mathbb{E}[(\bar{X} - \bar{Y})\mathrm{d}\bar{K}^{\bar{X}}] = 0$. This idea of projectors allows us to show uniqueness for the McKean–Vlasov SDE

(1.4) easily. This idea is also behind the proof of convergence of the particle system (1.2).

Concerning the convergence of the particle system (1.2), we use a pathwise approach introduced by Tanaka in [Tan84] and revisited in [CDFM20] (and further developed in [CL15, BCD20] in the rough path context): for any fixed $\omega \in \Omega$, the particle system (1.2) can also be viewed as a McKean–Vlasov equation (1.4), where, however, the law on the driving signal $\bar{W}$ is not the Wiener measure, but the random empirical measure $L^{W,N}(\omega) = \frac{1}{N}\sum_{i=1}^N \delta_{W^i(\omega)}$. Under this viewpoint, the term $\mathrm{d}K^N$ is also a projector in L^2 on the space of processes with average $q(t)$, but where the underlying measure is the empirical measure $L^{W,N}(\omega)$ instead of the Wiener measure. Intuitively then, since $L^{W,N}(\omega)$ converges P-a.s. to the Wiener measure, we expect that $\mathrm{d}K^N$ should converge to the projector under the Wiener measure, that is, $\mathrm{d}\bar{K}$; this should imply the convergence of the particle system to the McKean–Vlasov SDE (1.4).

However, a direct proof based only on this pathwise approach seems not easy. Indeed, to make this argument work, one needs to create an optimal coupling (in the sense of Wasserstein distance) between the Wiener measure and the random empirical measure $L^{W,N}(\omega)$, and such coupling does not have a Gaussian structure. Having a Gaussian measure on the driving signal allows us to use classical stochastic analysis tools like the Itô formula; such tools give in turn uniform BV estimates on $\mathrm{d}K^N$, which are also needed in the proof. Moreover, the pathwise argument gives a convergence only of the one-time marginals, that is, convergence of $\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}(\omega)}$ (as a random measure on $[0,1]$) to the law of $\bar{X}_t$ for every fixed $\bar{X}_t$, and not convergence of $\frac{1}{N}\sum_{i=1}^N \delta_{X_\cdot^{i,N}(\omega)}$ as a measure on the path space $C([0,T];[0,1])$.

For these reasons, we use first a tightness argument. Namely we show uniform (in N) BV and Hölder type bounds on the solution to the particle system (1.2), which give a tightness result; we show then that any limit point of (1.2) satisfies the McKean–Vlasov SDE (1.4), obtaining at once convergence of the particle system and existence for the McKean–Vlasov SDE itself. Once we have these uniform BV bounds and the existence for the McKean–Vlasov SDE, we can then use the pathwise argument explained before. From this pathwise argument, we also get the rate of convergence $O(1/\sqrt{\log(N)})$ for $\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}(\omega)}$.

The fact that $\mathrm{d}\bar{k}$ acts like a projector is classical and has been used since at least [LS84]. However, combining this fact with standard fixed-point arguments for McKean–Vlasov SDEs, as in [Szn91, Szn84], seems not easy in the presence of nonlocal effects of boundaries as here. The reason in our model is that $\mathbb{E}[\mathrm{d}\bar{k}]$ is just a BV and continuous term in time (not Lipschitz-continuous) without a distinguished sign. Even the use of the Lipschitz bounds from [DI91], on the sup norm of the reflecting term $\bar{k}$ in terms of the driving signal, seems not too helpful. For our model, the fact that $\mathrm{d}\bar{K}$ acts also as projector allows us to overcome this difficulty.

We remark that the action of $\mathrm{d}\bar{K}$ as projector on the processes with prescribed average, as well as some tricks used in the proof of uniform Hölder bounds for the particle system, are quite specific to our model. For extension to more general nonlocal effects of the reflecting boundary, other methods may be useful, which we do not explore here, for example, the penalization method (e.g., [Men83], used in [Jab17]), the approach based on Lions' derivative (e.g., [Lio08, CD18], used in [BCCdRH20]), pathwise approaches (e.g., [DGHT19, Aid16, FR13]), and PDE-based or singular interaction methods (see the paragraph below).

The PDE and singular interaction viewpoint. The Fokker–Planck equation (1.6) associated with the McKean–Vlasov SDE (1.4), that is, the equation for the probability density function (pdf) $u(t,\cdot)$ of $\bar{X}_t$, is a nonlinear nonlocal PDE. From this PDE (1.6), we can get another expression for $\mathrm{d}\bar{K}$ in terms of u:

$$\dot{\bar{K}}_t = \dot{q}(t) + \int_0^1 \mu(x)u(t,x)\mathrm{d}x + \frac{\sigma^2}{2}(u(1)-u(0)).$$

The nonlocal effects of the reflecting boundary appear as an additional drift term (here $u(1)-u(0)$) depending on the values of the pdf u at the boundary. Hence our model can be interpreted as a McKean–Vlasov SDE with reflecting boundaries and singular interaction at the boundary, and one may try to use PDE methods or an approach to singular interaction to deal with our model.

The literature about McKean–Vlasov SDEs with singular interaction and about a PDE-based approach to McKean–Vlasov SDEs is large, though we are not aware of a work which can easily cover our model. We only mention some references related to our model. We mention [Szn91, Chapter II], which deals with viscous Burgers equation as a McKean–Vlasov SDE with Dirac delta interaction (without boundaries), in particular the McKean–Vlasov SDE is also driven by the pdf of the solution. We also mention [BJ11, BJ15, BJ18], which study systems of interacting second-order diffusions (that is, SDEs for the acceleration of the particles) in a domain. Such systems contain a form of singular interaction and a form of reflection at the boundaries, though both interaction and reflection are of different types than in our model. Finally we mention [Kol07] for an approach based on semigroup theory to McKean–Vlasov SDEs and particle approximation.

In this paper we do not explore the PDE viewpoint and we give in section 2.2 a formal argument, without any rigorous proof, to show that (1.6) is indeed the Fokker–Planck equation for the SDE (1.4).

Organization of the paper. The paper is organized as follows. In section 2 we show the formal link, without rigorous proofs, between SDEs and Fokker–Planck equations, in presence of boundaries and mean-field interaction. In section 3 we give the precise setting and the main results. The proofs of these results are given in section 4. Finally, in the appendix we show well-posedness for the particle systems (1.2).

2. Review of PDEs and diffusion processes.

In this section we revisit the link between second-order PDEs and associated SDEs, both in presence of boundary and with nonlinearity, as the PDE (1.6) we consider here; we took inspiration from [Son07]. Our aim here is not to give rigorous results but to provide an easy yet clear "translator" between SDEs and PDEs, which applies, but is not restricted, to our case and shows in particular why (1.4) is the SDE corresponding to (1.6). For this reason, we keep all the computations at a formal level, without any rigorous proof.

In the following we focus our attention on the one-dimensional case, mostly for simplicity. We take $b : I \to \mathbb{R}$ a given vector field on $\mathbb{R}$ or on an interval I of $\mathbb{R}$ when specified, and $\sigma > 0$ positive constant; W is a real Brownian motion. For two functions $f, g : I \to \mathbb{R}$, we call

$$\langle f, g\rangle = \int_I f(x)g(x)\mathrm{d}x$$

their L^2 scalar product.

2.1. Diffusion, forward and backward PDEs.

It is well-known that the (forward) PDE

$$\partial_t p = (\sigma^2/2)\partial_y^2 p - \partial_y(bp) \text{ for } t > s,$$

with time-s initial data δ_x, models the evolution of the transition density function $p = p(s,x;t,y)$ of the diffusion process X, given by

$$\mathrm{d}X_t = \sigma \mathrm{d}W_t + b(t,X_t)\mathrm{d}t,\ X_s = x\ .$$

More generally, if u is the solution to this forward PDE with time-s initial data u_s, which is assumed to be a pdf, then u is the pdf of the solution of the same SDE but initial law $u_s(x)\mathrm{d}x$.

The dual viewpoint will be important. Consider the (backward) PDE

$$-\partial_s v = (\sigma^2/2)\partial_x^2 v + b\partial_x v$$

with terminal data $v(t,.) = \Psi$. Assuming v to be regular enough, Itô's formula gives a representation of the backward PDE solution v as follows:

$$v(s,x) = E[\Psi(X_t)|X_s = x]. \tag{2.1}$$

By a simple formal computation, one shows that

$$\frac{\mathrm{d}}{\mathrm{d}r}\langle p(s,x,r,\cdot), v(r,\cdot)\rangle = 0,$$

which implies

$$v(s,x) = \int \Psi(y)p(s,x,t,y)\mathrm{d}y. \tag{2.2}$$

At last, comparing (2.1) and (2.2) shows that $p(s,x,t,y)\mathrm{d}y$ is indeed the law of X_t, started at $X_s = x$, as claimed in the beginning of this subsection.

2.2. Reflected diffusion and PDEs with boundary. We now discuss the case of a spatial domain, with a focus on the simple case $I = [0,1]$. A reflected SDE on the domain $[0,1]$ is an SDE of the form

$$\begin{aligned}
&\mathrm{d}X_t^\circ = \sigma \mathrm{d}W_t + b(X_t^\circ)\mathrm{d}t - \mathrm{d}k_t\ ,\ X_s^\circ = x,\\
&X_t^\circ \in \bar{I}\ \ \forall t \geq s,\\
&\mathrm{d}|k| = 1_{X_t^\circ \in \{0,1\}}\mathrm{d}|k|,\ \ \mathrm{d}k = n(X_t^\circ)\mathrm{d}|k|.
\end{aligned}$$

Here, $n(0) = -1, n(1) = +1$ are the outer normals of our domain $[0,1]$. The solution is a couple (X,k) satisfying the above condition (it is implicitly assumed that k has BV paths). The last condition means that k acts only when X is on the boundary, giving a small "kick" so that X does not leave the domain I.

From the PDE viewpoint, the transition density function $p^\circ(s,x;t,y)$ associated with X is a solution to the following forward equation:

$$\begin{aligned}
&\partial_t p^\circ = (\sigma^2/2)\partial_y^2 p^\circ - \partial_y(bp^\circ) \text{ in } I^\circ \text{ for } t > s,\\
&(\sigma^2/2)\partial_y p^\circ - bp^\circ = 0 \text{ at } \partial I \text{ for } t > s,\\
&p^\circ(s,x,s,\cdot) = \delta_x,
\end{aligned}$$

More generally, if u is the solution to this forward PDE with time-s initial data u_s, assumed to be a pdf, then u is the pdf of the solution of the same SDE but initial law $u_s(x)\mathrm{d}x$. We show this fact formally using the dual viewpoint.

First step. Let v° be a regular solution to the dual backward equation, i.e., the time-t terminal value probelm with Neumann (no-flux) boundary data,

$$\begin{aligned} & -\partial_s v^\circ = (\sigma^2/2)\partial_x^2 v^\circ + b\partial_x v^\circ \text{ in } I^\circ \text{ for } s < t \ , \\ & \partial_x v^\circ(s,0) = \partial_x v^\circ(s,1) = 0 \text{ at } \partial I \ \forall \ s < t \ , \\ & v(t,.) = \Psi \end{aligned}$$

with some (regular) time-t terminal data Ψ. Then necessarily v° has the representation

$$v^\circ(s,x) = E[\Psi(X_t^\circ)|X_s^\circ = x]. \tag{2.3}$$

Indeed, the Itô formula gives

$$\begin{aligned} \mathrm{d}[v(s,X_s)] &= \partial_s v(X)\mathrm{d}s + \partial_x v(X)\mathrm{d}X + \frac{\sigma^2}{2}\partial_x^2 v(X)\mathrm{d}s \\ &= \left[\left(\partial_s + b\partial_x + \frac{\sigma^2}{2}\partial_x^2\right)v\right](X)\mathrm{d}s + \sigma\partial_x v(X)\mathrm{d}W - \partial_x v(X)\mathrm{d}k = \sigma\partial_x v(X)\mathrm{d}W, \end{aligned}$$

where we have used the equation for v to kill the first addend in the second line and the boundary conditions on v to kill the term with dk. Taking expectation, we get that $E[v(s,X_s)]$ is constant in s, which implies (2.3).

Second step. Again with simple formal computation, one shows

$$\frac{\mathrm{d}}{\mathrm{d}r}\langle p^\circ(s,x,r,\cdot), v^\circ(r,\cdot)\rangle = 0,$$

which implies

$$v^\circ(s,x) = \langle p_t, \Psi\rangle. \tag{2.4}$$

From (2.3) and (2.4) we conclude

$$\langle p^\circ(s,x,t,\cdot), \Psi\rangle = E[\Psi(X_t^\circ)|X_s^\circ = x] \ .$$

Since this is true for every regular Ψ, then p° is the law of X° conditional to $X_s^\circ = x$.

If u° is a solution to the same forward equation as above, but with generic initial condition u_s, integrating p in $u_s(x)$ shows that u is the pdf of the law of $X^{\circ,u}$, the process satisfying the same SDE but with initial law $u_s(x)\mathrm{d}x$.

2.3. McKean–Vlasov diffusion, nonlinear mean-field PDEs. Here we introduce the McKean–Vlasov setting. For a given drift $\bar{b} : \mathbb{R}\times\mathcal{P}(\mathbb{R}) \to \mathbb{R}$, where $\mathcal{P}(\mathbb{R})$ is the space of probability measures on $\mathbb{R}$, we consider

$$\mathrm{d}\bar{X}_t = \sigma\mathrm{d}W_t + \bar{b}(\bar{X}_t, \mathrm{Law}(\bar{X}_t))\mathrm{d}t, \quad \mathrm{Law}(\bar{X}_s) \text{ given.}$$

Existence and uniqueness for regular bounded $\bar{b}$ (regularity with respect to the measure variable is usually in terms of the Wasserstein distance) are proved via a fixed-point argument on the map $(m_t)_{t\geq s} \mapsto (\mathrm{Law}(X_t^m))_{t\geq s}$, where X^m is the solution to the (classical) SDE

$$\mathrm{d}X_t^m = \sigma\mathrm{d}W_t + \bar{b}[X_t^m, m_t]\mathrm{d}t, \quad \mathrm{Law}(X_s^m) = m_s.$$

See Sznitman [Szn91, Theorem 1.1] for the classical case of $\bar{b}$ linear in the measure argument or, for instance, [CDFM20] for a more general $\bar{b}$.

The corresponding forward Fokker–Planck equation now takes the form

$$\partial_t \bar{u} = (\sigma^2/2)\partial_y^2 \bar{u} - \partial_y[\bar{b}(y, \bar{u}_t(\cdot))\bar{u}] \tag{2.5}$$

with time-s initial data $\bar{u}_s = \mathrm{Law}(\bar{X}_s)$ (for notational reasons, we blur the difference between $u(t,.)$ and $u(t,y)dy$). Note that this is a nonlinear nonlocal PDE.

We again outline why $\bar{u}(t,.)$ is indeed the law of $\bar{X}_t$. Fix the family $\bar{u} := \{\bar{u}(t,.) : t \geq s\}$ and consider the (linear!) backward PDE

$$-\partial_s v = (\sigma^2/2)\partial_x^2 v + \bar{b}[x, \bar{u}(s,.)]\partial_x v$$

with terminal data $v_t = \Psi$. As in section 2.1, we see that the corresponding forward PDE (which by construction coincides with (2.5)) yields the law of a diffusion process $X^{\bar{u}}$, that is, $\mathrm{Law}(X_t^{\bar{u}}) = \bar{u}(t,\cdot)$. But then $X^{\bar{u}}$ solves the McKean–Vlasov SDE, therefore (by uniqueness for the McKean–Vlasov SDE) $X^{\bar{u}} = \bar{X}$ and so $\bar{u}_t$ is the pdf of the marginal $\bar{X}_t$ at time t.

2.4. Reflected McKean–Vlasov diffusion, mean-field PDE with boundary. We can consider the simplest case of reflected McKean–Vlasov SDE on I, as in [Szn84], namely

$$\begin{aligned}
&\mathrm{d}\bar{X}_t = \sigma \mathrm{d}W_t + \bar{b}[\bar{X}_t, \mathrm{Law}(\bar{X}_t)]\mathrm{d}t - \mathrm{d}\bar{k}_t, \quad \mathrm{Law}(\bar{X}_s) \text{ given},\\
&\bar{X}_t \in \bar{I} \quad \forall t \geq s,\\
&\mathrm{d}|\bar{k}| = 1_{\bar{X}_t \in \{0,1\}}\mathrm{d}|\bar{k}|, \quad \mathrm{d}\bar{k} = n(\bar{X}_t)\mathrm{d}|\bar{k}|.
\end{aligned}$$

Adapting the arguments in the two previous sections, one shows that the pdf of $\bar{X}_t$ is given via the following nonlinear nonlocal forward PDE:

$$\begin{aligned}
&\partial_t \bar{u}_t = (\sigma^2/2)\partial_y^2 \bar{u} - \partial_y[\bar{b}[y, \bar{u}_t(.)]\bar{u}] \text{ in } I^\circ \text{ for } t > s\ ,\\
&(\sigma^2/2)\partial_y \bar{u} - \bar{b}[\cdot, \bar{u}_t(.)]\bar{u} = 0 \text{ at } \partial I \text{ for } t > s,
\end{aligned}$$

with time-s initial data $\bar{u}_s = \mathrm{Law}(\bar{X}_s)$.

2.5. Interaction coming from the boundaries. As explained in the introduction, the SDE (1.4) does not fall in the previous class. Indeed, the drift depends not only on the law of $\bar{X}_t$ but also on the law of $\mathrm{d}\bar{k}_t$. We will not treat here the general case of drifts depending on the law of $\mathrm{d}\bar{k}$, but we focus our attention on our model.

We claim that the forward Fokker–Planck equation associated with (1.4) is (1.6): we show formally that if u satisfies the PDE (1.6), with initial condition u_0, then $u(t,x)$ is the pdf of the random variable $\bar{X}_t$ solving (1.4), with initial law u_0. The formal proof puts together the arguments for boundary problems and McKean–Vlasov SDEs with the additional difficulty of interaction coming from the boundary, for which we will use the constraint on the average in (1.6).

Take $\dot{\bar{K}}$ as in the PDE (1.6), and call $(X^{\bar{K}}, k^{\bar{K}})$ the solution of the reflecting SDE

$$\begin{aligned}
&\mathrm{d}X^{\bar{K}} = \sigma \mathrm{d}W_t + (-\mu(X^{\bar{K}}) + \dot{\bar{K}}(t))\mathrm{d}t - \mathrm{d}k^{\bar{K}}, \quad \mathrm{Law}(X^{\bar{K}}) = u_0,\\
&X_t^{\bar{K}} \in \bar{I} \quad \forall t \geq 0,\\
&\mathrm{d}|k^{\bar{K}}| = 1_{X_t^{\bar{K}} \in \{0,1\}}\mathrm{d}|k^{\bar{K}}|, \quad \mathrm{d}k^{\bar{K}} = n(X_t^{\bar{K}})\mathrm{d}|k^{\bar{K}}|.
\end{aligned}$$

As a consequence of section 2.2 (applied with given $\dot{\bar{K}}$), the law of $X_t^{\bar{K}}$ must be given by $u(t,\cdot)$, i.e., the PDE solution to (1.6). It remains to show that $X^\Lambda = \bar{X}$, the solution to the McKean–Vlasov (1.4). To this end, note that $\mathbb{E}[X_t^{\bar{K}}] = \int xu(t,x)\mathrm{d}x = q(t)$, using the basic constraint in (1.6). This shows that $X^{\bar{K}}$ is a solution to (1.4) and, by uniqueness of this equation (formally, see also Theorem 3.9), we conclude that $X^{\bar{K}} = \bar{X}$.

3. The setting and the main results.

3.1. The particle system. In the following, we consider a probability space $(\Omega, \mathcal{A}, P)$ and independent Brownian motions W^i, $i = 1 \dots N$, on a filtration $(\mathcal{F}_t)_t$ (satisfying the standard assumption). We are given a function of space $\mu : [0,1] \to \mathbb{R}$ and a function of time $q : [0,T] \to [0,1]$. The assumptions on μ and q will be given later. The noise intensity σ is assumed to be constant (possibly 0).

We consider the system of N interacting particles:

(3.1)
$$\begin{aligned}
\mathrm{d}X_t^{i,N} &= \left(-\mu(X_t^i) + \frac{1}{N}\sum_{j=1}^N \mu\left(X_t^{j,N}\right)\right)\mathrm{d}t + \dot{q}(t)\mathrm{d}t + \sigma\mathrm{d}W^i - \frac{1}{N}\sum_{j=1}^N \sigma\mathrm{d}W_t^j - \mathrm{d}k_t^{i,N} \\
&\quad + \frac{1}{N}\sum_{j=1}^N \mathrm{d}k_t^{j,N}, \quad i = 1, \dots N, \\
X^{i,N} &\in C([0,T];[0,1]), \ k^i \in C([0,T];\mathbb{R}) \text{ a.s.}, \quad i = 1, \dots N, \\
\mathrm{d}|k^{i,N}|_t &= 1_{X_t^{i,N} \in \{0,1\}}\mathrm{d}|k^{i,N}|_t, \quad \mathrm{d}k_t^{i,N} = n\left(X_t^{i,N}\right)\mathrm{d}|k^{i,N}|_t, \quad i = 1, \dots N.
\end{aligned}$$

Here n is the outer normal vector of the domain $]0,1[$ and $|k^{i,N}|$ is the total variation process associated with $k^{i,N}$ (note that it is not the modulus of $k^{i,N}$). A solution is a couple $(X^{(N)}, k^{(N)}) = (X^{i,N}, k^{i,N})_{i=1,\dots N}$ of an $(\mathcal{F}_t)_t$-progressively measurable continuous semimartingale $X^{(N)}$ and an $(\mathcal{F}_t)_t$-progressively measurable BV process $k^{(N)}$, satysfying the above system. Without loss of generality, we can assume $k_0^{(N)} = 0$. We will often omit the second superscript N (which denotes the number of particles) when not needed.

This system is the exact formulation of the interacting particle system (1.2), with the term $\mathrm{d}K^N$ in (1.2) given by

$$\mathrm{d}K^N = \dot{q}(t)\mathrm{d}t + \frac{1}{N}\sum_{j=1}^N \mu(X_t^j)\mathrm{d}t - \frac{1}{N}\sigma\sum_{j=1}^N \mathrm{d}W^j + \frac{1}{N}\sum_{j=1}^N \mathrm{d}k^j.$$

The term $-\mathrm{d}k^i$ represents the reflection at the boundary of $[0,1]$, while the term $\mathrm{d}K^N$ is independent of i and ensures that the empirical average $N^{-1}\sum_{i=1}^N X_t^i$ stays equal to $q(t)$.

The system (3.1) can be interpreted as an SDE for X on $[0,1]^N$ with oblique reflecting boundary conditions, where the direction of reflection keeps $X^{(N)}$ in the moving hyperplane $H_t = \{x \in \mathbb{R}^N \mid \frac{1}{N}\sum_{i=1}^N x^i = q(t)\}$. Another interpretation of this system is as an SDE on $H_t \cap [0,1]^N$ with normal boundary condition, where the domain, in the frame of H_t, is a moving convex polygon and the normal reflection is in the frame of H_t. This interpretation is used in the proof of well-posedness of the system (3.1) (Proposition 3.5).

We work under the following assumptions on μ and q (and $X_0^{(N)}$).

CONDITION 3.1. (i) *The function* $-\mu$ *is* C^2 *on* $]0,1[$ *and one-side Lipschitz-continuous, namely, there exists* $c \geq 0$ *such that, for every* x, y *in* $]0,1[$,

$$-(\mu(x)-\mu(y))(x-y) \leq c|x-y|^2 \quad \forall x,y \in]0,1[. \tag{3.2}$$

(ii) *The function* μ *satisfies*

$$\sup_{x\in]0,1/2[} |x||\mu(x)| + \sup_{x\in]1/2,1[} |1-x||\mu(x)| < +\infty.$$

(iii) *There exists* $0<\rho<1/2$ *such that*

$$\mathit{sign}(x-1/2)\mu(x) \geq 0 \quad \forall x \in]0,\rho[\cup]1-\rho,\rho[; \tag{3.3}$$

moreover $\mu(0)=\mu(1)=0$.

CONDITION 3.2. *The map* q *is a Lipschitz-continuous function of time (in particular* $\dot{q}(t)$ *exists for a.e.* t*) and there exists* $0<\xi<1$ *such that* $\xi \leq q(t) \leq 1-\xi$ *for every* t.

CONDITION 3.3. *The (possibly random) initial datum* $X_0^{(N)}$ *is* $\mathcal{F}_0$*-measurable and verifies* $0 \leq X_0^{i,N} \leq 1$ *for* $i=1,\dots N$ *and* $\frac{1}{N}\sum_{i=1}^N X_0^{i,N} = q(0)$ P*-a.s.*

The typical example we have in mind for μ is the derivative of a double-well potential, with logarithmic divergence at the boundary, and the typical example for q is a piecewise linear continuous function which does not touch 0 or 1. These examples are used in the battery model from [GGM$^+$18].

While Condition 3.1(i) on μ and Condition 3.2 on q are structural assumptions of our model, Condition 3.1(ii) seems not really necessary: if, for example, μ diverges like $1/x^\alpha$ for some $\alpha>1$, we would expect that the system does not even touch the boundary, hence the classical McKean–Vlasov approach should apply, but for technical reasons our proof does not apply to this situation; see Remark 4.11. Condition 3.1(iii) is also technical and we expect that it can be removed without too much effort; see Remark 4.24.

Actually we do not work directly with the system (3.1) but, to avoid possible singularity of μ at the boundary, we take a regularization μ^ϵ, C^2 on the closed domain $[0,1]$, with $\mu^\epsilon=\mu$ on $[\epsilon,1-\epsilon]$ and $|\mu^\epsilon| \leq |\mu|$ on $]0,1[$ and verifying both the one-side Lipschitz condition (3.2) and the condition (3.3) uniformly in ϵ. We then consider the system

(3.4)
$$\begin{aligned}
\mathrm{d}X_t^i &= \left(-\mu^\epsilon(X_t^i) + \frac{1}{N}\sum_{j=1}^N \mu^\epsilon(X_t^j)\right)\mathrm{d}t + \dot{q}(t)\mathrm{d}t + \sigma \mathrm{d}W_t^i - \frac{1}{N}\sum_{j=1}^N \sigma \mathrm{d}W_t^j - \mathrm{d}k_t^i \\
&\quad + \frac{1}{N}\sum_{j=1}^N \mathrm{d}k_t^i, \quad i=1,\dots N, \\
&X^i \in C([0,T];[0,1]),\ k^i \in C([0,T];\mathbb{R}) \text{ a.s.}, \quad i=1,\dots N, \\
&\mathrm{d}|k^i|_t = 1_{X_t^i\in\{0,1\}}\mathrm{d}|k^i|_t, \quad \mathrm{d}k_t^i = n(X_t^i)\mathrm{d}|k^i|_t, \quad i=1,\dots N.
\end{aligned}$$

When we want to stress the dependence on N and ϵ, we write $X^{i,N,\epsilon}$ and $X^{(N,\epsilon)} = (X^1,\dots X^N)$ and similarly for $k^{i,N,\epsilon}$, $k^{(N,\epsilon)}$.

Remark 3.4. Here and in the following, when we talk about pathwise uniqueness, resp., uniqueness in law, we refer to pathwise uniqueness of $X^{(N,\epsilon)}$, resp., of the law of $X^{(N,\epsilon)}$. Uniqueness of $X^{(N,\epsilon)}$ implies in turn uniqueness of $k^{(N,\epsilon)} - \mathbb{E}k^{(N,\epsilon)}$, but we do not make any uniqueness statement on $k^{(N,\epsilon)}$ itself.

PROPOSITION 3.5. *Assume Conditions* 3.2 *and* 3.3 *and assume that* μ^ϵ *is Lipschitz-continuous on* $[0,1]$*. Then there exists a solution to the particle system* (3.4) *and this solution is pathwise unique in* $X^{(N,\epsilon)}$*.*

The basic idea of the proof is simple: namely the SDE above is an SDE on the moving domain $H_t \cap [0,1]^N$ with normal boundary conditions. However, the proof is slightly technical and postponed to the appendix.

Remark 3.6. Similarly to (3.1), any solution to (3.4) satisfies $\frac{1}{N}\sum_{i=1}^N X_t^i = q(t)$ for every t.

3.2. The McKean–Vlasov SDE. In the following, we consider again a probability space $(\Omega, \mathcal{A}, P)$ and a Brownian motion W on a filtration $(\mathcal{F}_t)_t$ (satisfying the standard assumption); $\mathbb{E}$ denotes the expectation with respect to P. The functions μ, q, and σ are as in the previous subsection.

We consider the McKean–Vlasov SDE

(3.5)
$$\begin{aligned}
&\mathrm{d}\bar{X}_t = -\mu(\bar{X}_t)\mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}\bar{K}_t - \mathrm{d}\bar{k}_t,\\
&\int_0^T \mathbb{E}[|\mu(\bar{X}_r)|]\mathrm{d}r < \infty, \;\; \mathbb{E}\int_0^T \mathrm{d}|\bar{k}|_r < \infty, \;\; \mathrm{d}\bar{K}_t = (\mathbb{E}[\mu(\bar{X}_t)] + \dot{q}(t))\mathrm{d}t + \mathbb{E}[\mathrm{d}\bar{k}_t],\\
&\bar{X} \in C([0,T];[0,1]), \; \bar{k} \in C([0,T];\mathbb{R}) \text{ a.s.},\\
&\mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t \in \{0,1\}}\mathrm{d}|\bar{k}|_t, \;\; \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t.
\end{aligned}$$

Here again n is the outer normal vector of the domain $]0,1[$ and $|\bar{k}|$ is the total variation process associated with $\bar{k}$ (not the modulus of $\bar{k}$). A solution is a couple $(\bar{X}, \bar{k})$ of an $(\mathcal{F}_t)_t$-progressively measurable continuous semimartingale $\bar{X}$ and an $(\mathcal{F}_t)_t$-progressively measurable BV process $\bar{k}$, satisfying the above equation. Without loss of generality, we can assume $\bar{k}_0 = 0$. We sometimes say that $\bar{X}$ is a solution if there exists a process $\bar{k}$ such that $(\bar{X}, \bar{k})$ is a solution.

The assumptions on μ and q remain unchanged with respect to the particle system. In the assumption on $\bar{X}_0$, here the empirical average is replaced by the average with respect to P.

CONDITION 3.7. *The (possibly random) initial datum* $\bar{X}_0$ *is* $\mathcal{F}_0$*-measurable and verifies* $0 \le \bar{X}_0 \le 1$ *and* $\mathbb{E}\bar{X}_0 = q(0)$*.*

Remark 3.8. As for the particle system, it is easy to see that (under Condition 3.7) any solution to (3.5) satisfies $\mathbb{E}\bar{X}_t = q(t)$ for every t.

In view of the proof of convergence of the particle system, it is convenient to write (3.5) in the following equivalent way:

$$\begin{aligned}
&\mathrm{d}\bar{X}_t = \dot{q}(t)\mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}\bar{Z}_t - \mathbb{E}\mathrm{d}\bar{Z}_t,\\
&\int_0^T \mathbb{E}[|\mu(\bar{X}_r)|]\mathrm{d}r < \infty, \;\; \mathbb{E}\int_0^T \mathrm{d}|\bar{k}|_r < \infty, \;\; \mathrm{d}\bar{Z}_t = -\mu(\bar{X}_t)\mathrm{d}t - \mathrm{d}\bar{k}_t,\\
&\bar{X} \in C([0,T];[0,1]), \; \bar{k} \in C([0,T];\mathbb{R}) \text{ a.s.},\\
&\mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t \in \{0,1\}}\mathrm{d}|\bar{k}|_t, \;\; \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t.
\end{aligned} \tag{3.6}$$

3.3. The main results. Our main results are well-posedness of the McKean–Vlasov SDE (3.5) and convergence of the particle system (3.4) to the McKean–Vlasov SDE as $N \to \infty$ and $\epsilon \to 0$.

THEOREM 3.9. *Take a probability space* $(\Omega, \mathcal{A}, P)$*, a Brownian motion* W *on a filtration* $(\mathcal{F}_t)_t$ *(satistying the standard assumption), and an initial condition* $\bar{X}_0$*, assume Conditions* 3.1, 3.2*, and* 3.7*. Then there exists a unique solution* $(\bar{X}, \bar{k})$ *to the McKean–Vlasov SDE* (3.5)*.*

Let $\bar{X}$ be the solution to the McKean–Vlasov SDE (3.5) with initial datum $\bar{X}_0$ and, for $\epsilon > 0$, N in $\mathbb{N}$, let $(X^{1,N,\epsilon}, \dots X^{N,N,\epsilon})$ be the solution to the particle system (3.4) with initial datum $(X_0^1, \dots X_0^N)$. For E Polish space and Y^i E-valued random variables, we consider the empirical measures $\frac{1}{N}\sum_{i=1}^N Y^i$ as a $\mathcal{P}(E)$-valued random variable, where $\mathcal{P}(E)$ is the space of probability measures on E, endowed with the Borel σ-algebra with respect to the weak convergence (convergence against $C_b(E)$ test functions).

THEOREM 3.10. *Assume Conditions* 3.1 *and* 3.2*, Condition* 3.7 *on* $\bar{X}_0$*, and Condition* 3.3 *on* $(X_0^1, \dots X_0^N)$*. Assume also that* $\frac{1}{N}\sum_{i=1}^N \delta_{X_0^{i,N}}$ *converges in probability to Law*$(\bar{X}_0)$ *as* $N \to \infty$*. Then the sequence of empirical measures* $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N,\epsilon}}$ *on* $\mathcal{P}(C([0,T]))$ *converges in probability to Law*$(\bar{X})$*, as* $\epsilon \to 0$ *and* $N \to \infty$*.*

Note that the convergence result of the particle system (3.4) holds as $N \to \infty$, $\epsilon \to 0$ with no further restriction. In particular, one could send first ϵ to 0 and then N to ∞ to show the convergence of the original particle system (3.1).

Remark 3.11. The assumptions on the initial conditions may sound a bit rigid, in particular they cannot be satisfied taking $(X_0^1, \dots X_0^N)$ independent and identically distributed (i.i.d.) copies of $\bar{X}_0$ (the empirical average is not $q(0)$ for a.e. ω). However,

- an easy example of $(X_0^1, \dots X_0^N)$ satisfying this constraint is given by taking Y^i i.i.d. copies of a variable $\bar{X}_0$ with mean $q(0)$ and $X_0^i = Y^i - \frac{1}{N}\sum_{j=1}^N Y^j + q(0)$: by the law of large number $\frac{1}{N}\sum_{j=1}^N Y^j$ tends to $q(0) = \mathbb{E}\bar{X}_0$ and so the empirical measure of X_0^i tends to the law of $\bar{X}_0$ in probability;
- the assumptions can easily be relaxed allowing $q(0) = q^N(0)$ to be random and dependent on N, but keeping deterministic increments $q^N(t) - q^N(0)$, with $q^N(0)$ tending to $q(0)$ in probability as $N \to \infty$. This allows us to include the case of $(X_0^1, \dots X_0^N)$ i.i.d. copies of $\bar{X}_0$.

Finally, we give another convergence result and exhibit a rate of convergence for the time marginals. We denote by $\mathcal{W}_{2,[0,1]}$ the 2-Wasserstein distance on $[0,1]$.

PROPOSITION 3.12. *Assume that* μ *is* C^2 *on* $[0,1]$ *so that we can take* $\mu^\epsilon = \mu$ *for every* $\epsilon > 0$*. Assume the conditions of Theorem* 3.10 *and assume also that* $X_0^{i,N} = Y^i + \sum_{j=1}^N Y^j + q(0)$*, where* $(Y^i)_{i\in\mathbb{N}}$ *is a sequence of i.i.d. random variables with law Law*$(\bar{X}_0)$ *(see Remark* 3.11*). Then we have the following rate of convergence:*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \mathcal{W}_{2,[0,1]}\left(Law(\bar{X}_t), \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}\right)\right] = O(1/\sqrt{\log(N)}) \qquad \text{as } N \to \infty.$$

4. The proof.

4.1. The strategy. The strategy of the proof is as follows:

- We first prove uniqueness for the McKean–Vlasov SDE. For later use, we prove uniqueness among a larger class of solutions, namely possibly non-adapted processes. We also give a stability result with respect to the drift μ.

- For convergence of the particle system and existence of the McKean–Vlasov SDE, we prove uniform (in N and ϵ) BV and Hölder estimates for $k^{i,N,\epsilon}$ and uniform Hölder estimates for $X^{i,N,\epsilon}$. These estimates in turn imply tightness for the empirical measures $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N,\epsilon}}$ and more generally for $\frac{1}{N}\sum_{i=1}^N \delta_{(W^i,X^{i,N,\epsilon},-\int_0^\cdot \mu^\epsilon(X_r^{i,N,\epsilon})\mathrm{d}r-k^{i,N,\epsilon})}$.
- We then prove that any limit point of the empirical measures $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N,\epsilon}}$ is the law of a (possibly nonadapted) solution to the McKean–Vlasov SDE. Uniqueness of the McKean–Vlasov SDE implies that the whole sequence of empirical measure converges to the law of the unique solution and that this solution is actually adapted.
- Finally, we prove the rate of convergence using a pathwise approach. We first show that particle system (3.4) can be interpreted as the McKean–Vlasov equation with a different measure on the inputs. The core of the proof is then a stability result of the McKean–Vlasov equation with respect to the inputs.

In the following subsections, we will use the letter C to denote a positive constant, whose value may change from line to line; we will sometimes use C_p to stress the dependence on p.

Remark 4.1. Here we comment about a possible alternative strategy, taken from [BEH18, BCdRGL20]. One could try to apply the penalization method used in those works to (1.4), where in this case the penalized equation is a reflected equation. Reducing the problem to its bare bones and in order to make it as similar as possible to [BEH18, BCdRGL20], we can look at the following equation:

$$\begin{aligned} &\mathrm{d}\bar{X}_t = \mathrm{d}W_t + \mathrm{d}\bar{K}_t - \mathrm{d}\bar{k}_t,\\ &\bar{X}_t \le 1, \quad \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t, \quad \mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t=1}\mathrm{d}|\bar{k}|_t,\\ &\mathbb{E}\bar{X}_t \ge \frac{1}{2}, \quad \bar{K} \text{ deterministic.} \end{aligned} \tag{4.1}$$

The penalized version of (4.1) is the following:

$$\begin{aligned} &\mathrm{d}\bar{X}_t^n = \mathrm{d}W_t + \mathrm{d}\bar{K}_t^n - \mathrm{d}\bar{k}_t^n,\\ &\bar{X}_t \le 1, \quad \mathrm{d}\bar{k}_t = n(\bar{X}_t)\mathrm{d}|\bar{k}|_t, \quad \mathrm{d}|\bar{k}|_t = 1_{\bar{X}_t=1}\mathrm{d}|\bar{k}|_t,\\ &\bar{K}_t^n := \int_0^t \varphi_n\left(\mathbb{E}[\bar{X}_s^n] - \frac{1}{2}\right)\mathrm{d}s, \end{aligned} \tag{4.2}$$

where $\varphi_n(x) = r1_{x\le -\frac{1}{n}} - nrx1_{-\frac{1}{n}<x\le 0}$ and $r>0$ is to be choose accordingly. Equation (4.2) is well-posed for every n because of [Szn84].

The goal is now to construct a solution to (4.1) as a limit, for $n\to\infty$, of a sequence of solutions $\bar{X}^n$ to (4.2). When proving that $\bar{X}^n$ is a Cauchy sequence in L^2 one gets

$$\mathbb{E}\left[|\bar{X}_t^n - \bar{X}_t^m|^2\right] \le -2\int_0^t \mathbb{E}[\bar{X}_s^n]\varphi_m\left(\mathbb{E}[\bar{X}_s^m] - \frac{1}{2}\right)ds - 2\int_0^t \mathbb{E}[\bar{X}_s^m]\varphi_n\left(\mathbb{E}[\bar{X}_s^n] - \frac{1}{2}\right)ds.$$

Since φ is bounded and nonnegative, one could conclude by proving that $\mathbb{E}[\bar{X}_s^n] - \frac{1}{2} \ge -\frac{c}{n}$ for some constant $c>0$. By taking the expectation in (4.2) we get

$$\mathbb{E}[\bar{X}_t^n] = \mathbb{E}[\bar{X}_s^n] + \int_s^t \varphi_n\left(\mathbb{E}[\bar{X}_u^n] - \frac{1}{2}\right)\mathrm{d}u - \mathbb{E}[\bar{k}_t^n - \bar{k}_s^n].$$

At this point we meet an additional difficulty with respect to [BCCdRH20]: in order to conclude as in the argument in [BCCdRH20], one would need $\mathbb{E}[k_\cdot]$ to be a Lipschitz function of time; however, $k_\cdot$ is in general only of bounded variation.

Maybe one could try to penalize both reflection terms. However, this is beyond the scope of the present paper.

4.2. Uniqueness and stability. In this subsection we establish uniqueness and stability results for the McKean–Vlasov SDE (3.5). The following result proves the uniqueness part of Theorem 3.9.

PROPOSITION 4.2. *Assume Condition* 3.1(i) *on* μ *and that* q *is measurable bounded (Conditions* 3.1 *and* 3.2 *in particular are enough). Assume also Condition* 3.7 *on* $\bar{X}_0$. *Strong uniqueness holds for the McKean–Vlasov SDE* (3.5). *Moreover, if* $\bar{X}$ *and* $\bar{Y}$ *are two solution to* (3.5) *starting from* $\bar{X}_0$, $\bar{Y}_0$, *with* $\mathbb{E}[\bar{X}_0] = \mathbb{E}[\bar{Y}_0] = q(0)$, *it holds for some* $C > 0$ *(independent of* $\bar{X}_0$ *and* $\bar{Y}_0$*), for every* t,

$$\mathbb{E}|\bar{X}_t - \bar{Y}_t|^2 \leq e^{2Ct}\mathbb{E}|\bar{X}_0 - \bar{Y}_0|^2.$$

Proof. It is enough to prove stability. We will use the superscripts $\bar{X}$, $\bar{Y}$ for the quantities $\bar{K}$, $\bar{k}, \ldots$ associated with $\bar{X}$, $\bar{Y}$. By the Itô formula for continuous semimartingales [RY99] we have

$$\begin{aligned} &\mathrm{d}|\bar{X} - \bar{Y}|^2 \\ &\quad = 2(\bar{X} - \bar{Y})(-\mu(\bar{X}) + \mu(\bar{Y}))\mathrm{d}t + 2(\bar{X} - \bar{Y})\mathrm{d}\bar{K}^{\bar{X}} - 2(\bar{X} - \bar{Y})\mathrm{d}\bar{K}^{\bar{Y}} \\ &\qquad - 2(\bar{X} - \bar{Y})\mathrm{d}\bar{k}^{\bar{X}} + 2(\bar{X} - \bar{Y})\mathrm{d}\bar{k}^{\bar{Y}}. \end{aligned}$$

For the first addend, the one-side Lipschitz condition of μ implies

$$(\bar{X} - \bar{Y})(-\mu(\bar{X}) + \mu(\bar{Y})) \leq c|\bar{X} - \bar{Y}|^2.$$

For the addends with $\bar{k}$, the orientation of $\bar{k}$ (as the outward normal) implies

$$-\int_0^t (\bar{X} - \bar{Y})\mathrm{d}\bar{k}^{\bar{X}} \leq 0$$

and similarly for $(\bar{X} - \bar{Y})\mathrm{d}\bar{k}^{\bar{Y}}$. For the addends with K, we take the expectation and use that K is deterministic and that $E[\bar{X}_t] = E[\bar{Y}_t] = q(t)$ (see Remark 3.8): we obtain

$$\mathbb{E}\int_0^t (\bar{X} - \bar{Y})\mathrm{d}K^{\bar{X}} = \int_0^t (\mathbb{E}[\bar{X}] - \mathbb{E}[\bar{Y}])\mathrm{d}K^{\bar{X}} = 0.$$

Putting it all together, we get

$$\mathbb{E}|\bar{X}_t - \bar{Y}_t|^2 \leq \mathbb{E}|\bar{X}_0 - \bar{Y}_0|^2 + C\int_0^t \mathbb{E}|\bar{X}_r - \bar{Y}_r|^2\mathrm{d}r.$$

We conclude by the Gronwall inequality. ∎

PROPOSITION 4.3. *Assume Condition* 3.1(i) *and that* q *is measurable bounded. Let* μ^n *a sequence of functions, with uniformly bounded one-side Lipschitz constant, converging uniformly to* μ *on every compact subset of* $]0,1[$, *such that* $|\mu^n| \leq C|\mu|$ *on* $]0,1[$. *Call* $\bar{X}^n$, $\bar{X}$ *the solutions to the SDE* (3.5), *resp., with* μ^n, μ *and with the same initial condition. Then it holds, as* $n \to \infty$,

$$\sup_{t\in[0,T]} \mathbb{E}|\bar{X}_t - \bar{X}^n_t|^2 \to 0.$$

Proof. By the Itô formula we have, proceeding as in the previous proof,

$$\mathrm{d}|\bar X-\bar X^n|^2 = 2(\bar X-\bar X^n)(-\mu(\bar X)+\mu^n(\bar X))\mathrm{d}t + 2(\bar X-\bar X^n)(-\mu^n(\bar X) + \mu^n(\bar X^n))\mathrm{d}t + \mathrm{d}(\text{other terms}),$$

where the other terms have nonpositive expectation. For the second addend, the uniform one-side Lipschitz condition implies, for some $c>0$ independent of n,

$$(\bar X-\bar X^n)(-\mu^n(\bar X)+\mu^n(\bar X^n)) \le c|\bar X-\bar X^n|^2.$$

For the first addend, the integrability condition on μ in (3.5) implies that, for every $\epsilon>0$, there exists $\delta>0$ such that

$$\mathbb{E}\int_0^T 1_{\bar X\notin[\delta,1-\delta]}|\mu(\bar X)|\mathrm{d}r<\epsilon$$

and similarly for μ^n since $|\mu^n|\le C|\mu|$. By the uniform convergence of μ^n to μ on $[\delta,1-\delta]$, there exists n_0 such that, for every $n\ge n_0$, $|\mu^n-\mu|<\epsilon$ on $[\delta,1-\delta]$. Therefore we have

$$\begin{aligned}&\mathbb{E}\int_0^t (\bar X-\bar X^n)(-\mu(\bar X)+\mu^n(\bar X))\mathrm{d}r\\ &\quad\le \mathbb{E}\int_0^t 1_{\bar X\in[\delta,1-\delta]}|\mu(\bar X)-\mu^n(\bar X)|\mathrm{d}r + \mathbb{E}\int_0^t 1_{\bar X\notin[\delta,1-\delta]}(|\mu(\bar X)|+|\mu^n(\bar X)|)\mathrm{d}r \le C\epsilon.\end{aligned}$$

Finally we obtain, for every $n\ge n_0$,

$$\mathbb{E}|\bar X_t-\bar X^n_t|^2 \le C\epsilon + C\int_0^t \mathbb{E}|\bar X_r-\bar X^n_r|^2\mathrm{d}r.$$

We conclude again by the Gronwall lemma. ∎

For the proof of convergence of the particle system, it a slightly stronger uniqueness result is actually useful, among a generalized class of solutions. Given a probability space $(\Omega,\mathcal{A},P)$ and a Brownian motion W on it (with respect to its natural filtration), we call a generalized solution a couple $(\bar X,\bar k)$ of $\mathcal{A}\otimes\mathcal{B}([0,T])$-measurable maps, satisfying the system (3.5) (or equivalently (3.6)) P-a.s., without any adaptedness condition; we also do not require $\mathcal{A}$ to be complete with respect to P. We also call a weak generalized solution the object $(\Omega,\mathcal{A},W,\bar X,\bar k,P)$. Note that the system makes sense even without adaptedness, since the noise is additive. The difference with the usual concept of solution lies exactly in the lack of adaptability (and lack of completeness of the σ-algebra $\mathcal{A}$). We say that $\bar X$ is a generalized solution if there exists a $\mathcal{A}\otimes\mathcal{B}([0,T])$-measurable map $\bar k$ such that $(\bar X,\bar k)$ is a generalized solution.

LEMMA 4.4. *Assume Condition* 3.1(i) *and that q is measurable bounded. Assume also Condition* 3.7 *on $\bar X_0$. Given $(\Omega,\mathcal{A},P)$ and W, uniqueness holds among generalized solutions.*

Proof. Let $(\bar X,\bar k^{\bar X})$ and $(\bar Y,\bar k^{\bar Y})$ be two solutions. Then $\bar X-\bar Y$ is a BV and continuous process satisfying P-a.e.

$$\mathrm{d}(\bar X-\bar Y) = (-\mu(\bar X)+\mu(\bar Y))\mathrm{d}t + \mathrm{d}\left(\bar K^{\bar X}-\bar K^{\bar Y}\right) + \mathrm{d}\left(\bar k^{\bar X}-\bar k^{\bar Y}\right).$$

Each of the addends in the right-hand side (RHS) is BV and continuous; in particular we can fix ω (outside a P-null set in $\mathcal{A}$) and apply the chain rule to get the expression for the differential of $|\bar X-\bar Y|^2$. The rest of the proof goes as in the proof of Proposition 4.2. ∎

Another useful tool in view of particle convergence is the Yamada–Watanabe principle, which, roughly speaking, states that strong uniqueness and weak existence imply uniqueness in law and strong existence. Since we are working in a slightly nonstandard context, with McKean–Vlasov SDEs and with generalized solutions, we repeat the statements and the proofs for our case.

PROPOSITION 4.5 (Yamada–Watanabe, uniqueness in law). *Let* $(\Omega^i, \mathcal{A}^i, P^i)$, $i = 1, 2$, *be two probability spaces, with associated Brownian motions* W^i *and generalized solutions* $(\bar{X}^i, \bar{k}^i)$, $i = 1, 2$, *such that* $Law(\bar{X}_0^1) = Law(\bar{X}_0^2)$. *Then the laws of* $(W^1, \bar{X}^1)$ *and* $(W^2, \bar{X}^2)$ *coincide.*

Proof. We take $\hat{\Omega} = (C([0,T]) \times \mathbb{R}) \times C([0,T])^2 \times C([0,T])^2$, endowed with the Borel σ-algebra $\hat{\mathcal{A}} = \mathcal{B}(\hat{\Omega})$ (with respect to the uniform topology). We call $\hat{\omega} = ((w, x_0), (\gamma^1, \kappa^1), (\gamma^2, \kappa^2))$ a generic element of Ω. Let $P^{i,W^i,\bar{X}_0^i}$ be the conditional law of $(\bar{X}^i, \bar{k}^i)$ with respect to W^i and $\bar{X}_0^i$, $i = 1, 2$. We take on $\mathcal{B}(\hat{\Omega})$ the probability measure $\hat{P} = P^{W,\bar{X}_0} \otimes P^{1,w,x_0} \otimes P^{2,w,x_0}$, where $P^{W,\bar{X}_0}$ is the product of the Wiener measure and the law of $\bar{X}_0^1$. We define $\hat{W}(\hat{\omega}) = w$, $\hat{X}_0(\hat{\omega}) = x_0$, $(\hat{X}^i(\hat{\omega}), \hat{k}^i(\hat{\omega})) = (\gamma^i, \kappa^i)$, $i = 1, 2$, the canonical projections. Now, for each $i = 1, 2$, the law of $(\hat{W}, \hat{X}^i, \hat{k}^i)$ is the law of $(W^i, \bar{X}^i, \bar{k}^i)$, in particular $\bar{K}^{\hat{X}^i} = \bar{K}^{\bar{X}^i}$. Therefore $(\hat{X}^i, \hat{k}^i)$, $i = 1, 2$, are two generalized solutions to (3.5), defined on the same probability space $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{P})$ with respect to the same Brownian motion $\hat{W}$ and with the same initial datum $\hat{X}_0^1 = \hat{X}_0^2 = \hat{X}_0$ $\hat{P}$-a.s. By the uniqueness result, $\hat{X}^1$ and $\hat{X}^2$ must coincide $\hat{P}$-a.s. Hence (calling $\gamma_\#$ the push-forward of the projection on the γ component), $\gamma_\# P^{1,w,x_0}$ and $\gamma_\# P^{2,w,x_0}$, the conditional laws of $\hat{X}^1$ and $\hat{X}^2$ given $(\hat{W}, \hat{X}_0) = (w, x_0)$, coincide for $P^{W,\bar{X}_0}$-a.e. (w, x_0). Therefore $\text{Law}(W^1, \bar{X}_0^1, \bar{X}^1) = P^{W,\bar{X}_0} \otimes \gamma_\# P^{1,w,x_0}$ and $\text{Law}(W^2, \bar{X}_0^2, \bar{X}^2) = P^{W,\bar{X}_0} \otimes \gamma_\# P^{2,w,x_0}$ coincide. The proof is complete. ∎

PROPOSITION 4.6 (Yamada–Watanabe, strong existence). *The generalized solution* $(\bar{X}^1, \bar{k}^1)$ *is actually a strong solution to* (3.5)*, that is, it is progressively measurable with respect to* $(\mathcal{F}_t^{W^1,\bar{X}_0^1})_t$*, the filtration generated by* W^1*,* $\bar{X}_0^1$ *and the* P^1*-null sets (and similarly for* $(\bar{X}^2, \bar{k}^2)$*).*

Proof. We continue using the notation of the previous proof. Call $(\hat{\mathcal{F}}_t^{\hat{W},\hat{X}_0})_t$ the filtration generated by $\hat{W}$, $\hat{X}_0$ and the $P^{W,\bar{X}_0}$-null sets on $C([0,T]) \times \mathbb{R}$. Note that the conditional law of $(\hat{X}^1, \hat{k}^1, \hat{X}^2, \hat{k}^2$ given $(\hat{W}, \hat{X}_0) = (w, x_0)$ is $P^{1,w,x_0} \otimes P^{2,w,x_0}$. Hence, for $P^{W,\bar{X}_0}$-a.e. (w, x_0), conditioning to $(\hat{W}, \hat{X}_0) = (w, x_0)$, $\hat{X}^1$ and $\hat{X}^2$ coincide a.s. and are independent. Hence, for $P^{W,\bar{X}_0}$-a.e. (w, x_0) given, conditioning to $(\hat{W}, \hat{X}_0) = (w, x_0)$, $\hat{X}^1$ must coincide with an element $Y^T(w, x_0)$ a.s. The random element Y^T, extended on a $P^{W,\bar{X}_0}$-null set, defines a solution map $Y^T : C([0,T]) \times \mathbb{R} \to C([0,T])$ which is $\hat{\mathcal{F}}_T^{\hat{W},\hat{X}_0}$-measurable: indeed, for every Borel subset B of $C([0,T])$, $\{Y^T \in B\}$ coincides $P^{W,\bar{X}_0}$-a.s. with $\{(w, x_0) \mid \gamma_\# P^{1,w,x_0}(B) = 1\}$, which belongs to $\hat{\mathcal{F}}_T^{\hat{W},\hat{X}_0}$ (since $P^{1,w,x_0}(B)$ is $\hat{\mathcal{F}}_T^{\hat{W},\hat{X}_0}$-measurable). From the previous proof, we have

$$\text{Law}(W^1, \bar{X}_0^1, \bar{X}^1) = P^{W,\bar{X}_0} \otimes \gamma_\# P^{1,w,x_0} = \text{Law}(W, \bar{X}_0^1) \otimes \delta_{Y^T(w,x_0)},$$

therefore $\bar{X}^1 = Y^T(W^1, \bar{X}_0^1)$ P^1-a.s. Since $(W, \bar{X}_0^1)$ is measurable from $\mathcal{F}_T^{W^1,\bar{X}_0^1}$ to $\hat{\mathcal{F}}_T^{\hat{W},\hat{X}_0}$, we conclude that $\bar{X}^1$ is $\mathcal{F}_T^{W^1,\bar{X}_0^1}$-measurable.

Concerning progressive measurability, we can restrict W^1, $\bar{X}^1$, and $\bar{k}^1$ on $[0,t]$ and repeat the previous argument: calling $\pi^t : C([0,T]) \to C([0,t])$ the restriction operator, we get that $\pi_t(\bar{X}^1) = Y^t(\pi_t(W^1), \bar{X}_0^1)$ P^1-a.s. and $\pi_t(\bar{X}^1)$ is $\mathcal{F}_t^{W^1,\bar{X}_0^1}$-

measurable. Hence $\bar{X}$ is adapted and therefore progressively measurable, by continuity of its paths. Progressive measurability of $\bar{k}^1$ follows because, P^1-a.s.,

$$\mathrm{d}\bar{k}^1 = -\mathrm{d}\bar{X}^1 - \mu(\bar{X}^1)\mathrm{d}t + \mathrm{d}W^1 + \mathbb{E}^{P^1}[\mu(\bar{X}^1)]\mathrm{d}t + \mathbb{E}^{P^1}[\mathrm{d}\bar{k}^1].$$

The proof is complete. ∎

4.3. Compactness for the particle system. Here we consider the particle system (3.4) and we give estimates which are uniform in N and ϵ. We will often omit the superscripts N and ϵ for notational simplicity.

4.3.1. *BV* estimates. We start estimating the BV norm of the average of the drift over the particles. Throughout this subsection, we will assume Conditions 3.1(i) on μ, 3.2 on q, and 3.3 on X_0.

LEMMA 4.7. *For every $1 \le p < \infty$, it holds that*

$$\sup_{N,\epsilon} \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^T \left|\mu^\epsilon\left(X_r^{i,N,\epsilon}\right)\right| \mathrm{d}r\right)^p + \sup_{N,\epsilon} \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^T \left|\mathrm{d}k_r^{i,N,\epsilon}\right|\right)^p < +\infty.$$

The proofs of this lemma and of the next one use mainly two facts:

- the one-side Lipschitz property of $-\mu$ and the reflection condition on k^i, that is, k^i has the same sign of $n(X^i)$;
- the property $\frac{1}{N}\sum_{i=1}^N X_t^i - q(t) = 0$.

Let us explain briefly how these two facts yield the BV estimates. We will focus only on the bounds on k^i, the bounds on $\mu(X^i)$ being similar. As in the standard argument for boundary terms, we take the differential of $|X_t^i - q(t)|^2$:

$$\mathrm{d}|X^i - q|^2 = -2(X_i - q)\mathrm{d}k^i + 2(X^i - q)\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k^j + \dots. \tag{4.3}$$

Disgarding the interaction terms, using the reflection condition, we would get an inequality of the form

$$\mathrm{d}|X^i - q|^2 = -2(X_i - q)\mathrm{d}k^i \le -2|X_i - q|\mathrm{d}|k^i|.$$

This inequality would give a bound on $|X_i - q(t)|\mathrm{d}|k^i|$ and so on $\mathrm{d}|k^i|$ (since $|X_i - q(t)|$ is bounded from below when X^i is on the boundary). However, the interaction term $2(X^i - q)\frac{1}{N}\sum_{j=1}^N \mathrm{d}k^j$ in (4.3) cannot be bounded as before, due to the $+$ sign instead of the $-$ sign. To deal with it, first we average over i: thanks to the property $\frac{1}{N}\sum_{i=1}^N X_t^i - q(t) = 0$, the average of the interaction terms disappears:

$$\frac{1}{N}\sum_{i=1}^{N}(X^i - q)\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k^j = 0;$$

hence we get a bound on the average of $\mathrm{d}|k^i|$ (Lemma 4.7). Second, this bound allows us to control the interaction term $2(X^i - q)\frac{1}{N}\sum_{j=1}^N \mathrm{d}k^j$. Using this control in (4.3), we get a bound on $\mathrm{d}|k^i|$ for any i (Lemma 4.9).

Remark 4.8. The one-side Lipschitz condition on $-\mu$ and the regularity of μ in the interior $]0,1[$ imply that, for any $0 < c < 1/2$,

$$\sup_{\epsilon, x\in]0,1[} (\mu^\epsilon(x)\mathrm{sign}(x - 1/2))^- < +\infty,$$

$$\sup_{\epsilon, x\in [c,1-c]} |\mu^\epsilon(x)| < +\infty.$$

The condition $0 < \xi \leq q(t) \leq 1-\xi < 1$ implies that

$$(x-q(t))\mathrm{sign}(x-1/2)1_{x\notin[\xi/2,1-\xi/2]} \geq \frac{\xi}{2}1_{x\notin[\xi/2,1-\xi/2]}.$$

Putting together the above bounds, we get, for some $C \geq 0$ independent of x and ϵ, for every x in $]0,1[$ and every t,

$$\begin{aligned}
&(x-q(t))\mu^\epsilon(x)\\
&\quad = (x-q(t))\mathrm{sign}(x-1/2)(\mu^\epsilon(x)\mathrm{sign}(x-1/2))^+1_{x\notin[\xi/2,1-\xi/2]}\\
&\qquad - (x-q(t))\mathrm{sign}(x-1/2)(\mu^\epsilon(x)\mathrm{sign}(x-1/2))^-1_{x\notin[\xi/2,1-\xi/2]}\\
&\qquad - (x-q(t))\mu^\epsilon(x)1_{x\in[\xi/2,1-\xi/2]}\\
&\quad \geq \frac{\xi}{2}(\mu^\epsilon(x)\mathrm{sign}(x-1/2))^+1_{x\notin[\xi/2,1-\xi/2]} - C\\
&\quad = \frac{\xi}{2}|\mu^\epsilon(x)\mathrm{sign}(x-1/2)|1_{x\notin[\xi/2,1-\xi/2]} - \frac{\xi}{2}(\mu^\epsilon(x)\mathrm{sign}(x-1/2))^-1_{x\notin[\xi/2,1-\xi/2]} - C\\
&\quad \geq \frac{\xi}{2}|\mu^\epsilon(x)|1_{x\notin[\xi/2,1-\xi/2]} - C\\
&\quad \geq \frac{\xi}{2}|\mu^\epsilon(x)| - C.
\end{aligned}$$

By continuity of μ^ϵ on $[0,1]$, for $\epsilon > 0$ the same estimate holds on the closed interval $[0,1]$. From the reflection condition on k we also get

$$(X^i_t - q(t))\mathrm{d}k^i_t \geq \xi\mathrm{d}|k^i|.$$

Proof. By the Itô formula, we have

(4.4)

$$\begin{aligned}
&\mathrm{d}|X^i-q(t)|^2\\
&\quad = 2(X^i-q(t))\left(-\mu^\epsilon(X^i)+\frac{1}{N}\sum_{j=1}^N\mu^\epsilon(X^j)\right)\mathrm{d}t + 2\sigma(X^i-q(t))\left(\mathrm{d}W^i - \frac{1}{N}\sum_{j=1}^N\mathrm{d}W^j\right)\\
&\qquad + \sigma^2\left(1-\frac{1}{N}\right)\mathrm{d}t + 2(X^i-q(t))\left(-\mathrm{d}k^i+\frac{1}{N}\sum_{j=1}^N\mathrm{d}k^j\right).
\end{aligned}$$

We average over i. For the interaction term with $\frac{1}{N}\sum_{j=1}^N\mu^\epsilon(X^j)$, the condition $\frac{1}{N}\sum_{i=1}^N X^i_t = q(t)$ implies

$$\frac{1}{N}\sum_{i=1}^N(X^i-q(t))\frac{1}{N}\sum_{j=1}^N\mu^\epsilon(X^j)\mathrm{d}t = 0,$$

and similarly for the other interaction terms (with $\frac{1}{N}\sum_{j=1}^N\mathrm{d}W^j$ and with $\frac{1}{N}\sum_{j=1}^N\mathrm{d}k^j$). Hence we get

$$\begin{aligned}
&\mathrm{d}\frac{1}{N}\sum_{i=1}^N|X^i-q(t)|^2\\
&\quad = -2\frac{1}{N}\sum_{i=1}^N(X^i-q(t))\mu^\epsilon(X^i)\mathrm{d}t + 2\frac{1}{N}\sum_{i=1}^N\sigma(X^i-q(t))\mathrm{d}W^i\\
&\qquad + \sigma^2\left(1-\frac{1}{N}\right)\mathrm{d}t - 2\frac{1}{N}\sum_{i=1}^N(X^i-q(t))\mathrm{d}k^i.
\end{aligned}$$

Now we apply Remark 4.8 and obtain

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}|X_T^i-q(T)|^2+\xi\frac{1}{N}\sum_{i=1}^{N}\int_0^T|\mu^\epsilon(X^i)|\mathrm{d}r+2\xi\frac{1}{N}\sum_{i=1}^{N}\int_0^T\mathrm{d}|k^i|_r\\
&\quad\le\frac{1}{N}\sum_{i=1}^{N}|X_T^i-q(T)|^2+2\frac{1}{N}\sum_{i=1}^{N}\int_0^T(X^i-q(t))\mu^\epsilon(X^i)\mathrm{d}t+CT+2\frac{1}{N}\sum_{i=1}^{N}\int_0^T(X^i-q(t))\mathrm{d}k^i\\
&\quad\le\frac{1}{N}\sum_{i=1}^{N}|X_0^i-q(0)|^2+CT+2\sigma\left|\frac{1}{N}\sum_{i=1}^{N}\int_0^T(X_r^i-q(r))\mathrm{d}W_r^i\right|+\sigma^2T\\
&\quad\le C+2\sigma\left|\frac{1}{N}\sum_{i=1}^{N}\int_0^T(X_r^i-q(r))\mathrm{d}W_r^i\right|.
\end{aligned}
$$

By the Burkholder–Davis–Gundy inequality (and boundedness of X and q), we arrive at

$$
\begin{aligned}
&C'\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^T|\mu^\epsilon(X^i)|\mathrm{d}r\right)^p+C'\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^T\mathrm{d}|k^i|_r\right)^p\\
&\quad\le C+C\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^T|X_r^i-q(r)|^2\mathrm{d}r\right)^{p/2}\le C,
\end{aligned}
$$

which is the desired bound. ∎

Thanks to the previous lemma, we can conclude a uniform BV estimate on the drift.

LEMMA 4.9. *For every $1\le p<\infty$, it holds that*

$$
\sup_{N,\epsilon,i=1\ldots N}\mathbb{E}\left(\int_0^T\left|\mu^\epsilon\left(X^{\epsilon,N,i}\right)\right|\mathrm{d}r\right)^p+\sup_{N,\epsilon,i=1\ldots N}\mathbb{E}\left(\int_0^T\left|\mathrm{d}k_r^{N,\epsilon,i}\right|\right)^p<+\infty.
$$

Proof. We start as before from formula (4.4), for fixed i, and use Remark 4.8, getting

$$
\begin{aligned}
&|X_T^i-q(T)|^2+\xi\int_0^T|\mu^\epsilon(X^i)|\mathrm{d}r+2\xi\int_0^T\mathrm{d}|k^i|_r\\
&\quad\le|X_0^i-q(0)|^2+CT+2\int_0^T|X_r^i-q(r)|\frac{1}{N}\sum_{j=1}^{N}|\mu^\epsilon(X_r^j)|\mathrm{d}r\\
&\qquad+2\sigma\left|\int_0^T(X_r^i-q(r))\left(\mathrm{d}W^i-\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}W^j\right)\right|+\sigma^2T+2\int_0^T|X_r^i-q(r)|\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k_r^j\\
&\quad\le C+2\int_0^T\frac{1}{N}\sum_{j=1}^{N}|\mu^\epsilon(X_r^j)|\mathrm{d}r+2\sigma\left|\int_0^T(X_r^i-q(r))\left(\mathrm{d}W^i-\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}W^j\right)\right|+2\int_0^T\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k_r^j.
\end{aligned}
$$

By the Burkholder–Davis–Gundy inequality, we get

$$
\begin{aligned}
&C'\mathbb{E}\left(\int_0^T|\mu^\epsilon(X^i)|\mathrm{d}r\right)^p+C'\mathbb{E}\left(\int_0^T\mathrm{d}|k^i|_r\right)^p\\
&\quad\le C+C\mathbb{E}\left(\int_0^T\frac{1}{N}\sum_{j=1}^{N}|\mu^\epsilon(X_r^j)|\mathrm{d}r\right)^p+C+C\mathbb{E}\left(\int_0^T\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k_r^j\right)^p.
\end{aligned}
$$

Here we use Lemma 4.7 and conclude

$$C'\mathbb{E}\left(\int_0^T |\mu^\epsilon(X^i)|\mathrm{d}r\right)^p + C'\mathbb{E}\left(\int_0^T \mathrm{d}|k^i|_r\right)^p \le C.$$

The proof is complete. ∎

4.3.2. Hölder estimates. In this paragraph we use a similar strategy to estimate the Hölder norm of X^i, first taking the average over i to remove the interaction term, then using the estimate on the average to control the interaction term. In order to bound the Hölder norm, we take the Itô differential of $|X_t^i - q(t) - X_s^i + q(s)|^2$ (instead of just $|X_t^i - q(t)|^2$).

Throughout this subsection, we will assume Conditions 3.1(i), (ii) on μ, 3.2 on q, and 3.3 on X_0.

We start with a preliminary result which will be used in the next estimates.

LEMMA 4.10. *For every $1 \le p < \infty$, it holds for some $C_p \ge 0$ independent of s, t,*

$$\sup_{N,\epsilon,i} \mathbb{E}\left(\int_s^t \left[(X_r^i - q(r) - X_s^i + q(s))\mu^\epsilon(X_r^i)\right]^- \mathrm{d}r\right)^p \le C_p|t-s|^p,$$

$$\sup_{N,\epsilon,i} \mathbb{E}\left(\int_s^t \left[(X_r^i - q(r) - X_s^i + q(s))n(X_r^i)\right]^- |\mathrm{d}k_r^i|\right)^p \le C_p|t-s|^p.$$

Proof. We start with the first inequality and we fix $\delta > 0$ small, independently of ϵ and N. Using the elementary inequality $[a+b]^- \le |a| + [b]^-$ for $a, b \in \mathbb{R}$, we have

$$\begin{aligned}
&[(X_r^i - q(r) - X_s^i + q(s))\mu^\epsilon(X_r^i)]^- \\
&\quad \le |q(r) - q(s)||\mu^\epsilon(X_r^i)| + [(X_r^i - X_s^i)\mu^\epsilon(X_r^i)]^- \\
&\quad \le |q(r) - q(s)||\mu^\epsilon(X_r^i)| + |X_r^i - X_s^i| \max_{[\delta,1-\delta]} |\mu^\epsilon| \\
&\qquad + \left(1_{X_r^i<\delta, X_r^i\le X_s^i} + 1_{1-\delta<X_r^i, X_s^i\le X_r^i}\right)|X_r^i - X_s^i|[\mathrm{sign}(X_r^i - 1/2)\mu^\epsilon(X_r^i)]^- \\
&\qquad + \left(1_{X_s^i<X_r^i<\delta} + 1_{1-\delta<X_r^i<X_s^i}\right)|X_r^i - X_s^i||\mu^\epsilon(X_r^i)|.
\end{aligned}$$

For the first addend in the RHS, the Lipschitz property of q and Lemma 4.9 give

$$\mathbb{E}\left(\int_s^t |q(r) - q(s)||\mu^\epsilon(X_r^i)|\mathrm{d}r\right)^p \le C|t-s|^p \mathbb{E}\left(\int_0^T |\mu^\epsilon(X_r^i)|\mathrm{d}r\right)^p \le C_p|t-s|^p.$$

For the second and third addends, we have by Remark 4.8 (recall δ is fixed and X^i is in $[0, 1]$)

$$\mathbb{E}\left(\int_s^t |X_r^i - X_s^i| \max_{[\delta,1-\delta]} |\mu^\epsilon|\mathrm{d}r\right)^p \le C_p|t-s|^p,$$

$$\mathbb{E}\left(\int_s^t (1_{X_r^i<\delta, X_r^i\le X_s^i} + 1_{1-\delta<X_r^i, X_s^i\le X_r^i})|X_r^i - X_s^i|[\mathrm{sign}(X_r^i - 1/2)\mu^\epsilon(X_r^i)]^- \mathrm{d}r\right)^p \le C_p|t-s|^p.$$

Concerning the fourth addend, we consider only the case $1 - \delta < X_r^i < X_s^i$, the case $X_s^i < X_r^i < \delta$ being completely analogous. By the assumption of Condition 3.1(ii) we have

$$1_{1-\delta<X_r^i<X_s^i}|X_r^i - X_s^i||\mu^\epsilon(X_r^i)| \le C 1_{1-\delta<X_r^i<X_s^i} \sup_{1-\delta<x<1} (1-x)|\mu(x)| \le C.$$

Therefore, reasoning similarly for $X^i_s < X^i_r < \delta$, we have

$$\mathbb{E}\left(\int_s^t (1_{1-\delta<X^i_r<X^i_s} + 1_{1-\delta<X^i_r<X^i_s})|X^i_r - X^i_s||\mu^\epsilon(X^i_r)|\mathrm{d}r\right)^p \le C_p|t-s|^p.$$

Putting it all together, we obtain the first estimate.

For the second estimate, recall that $(X^i_r - X^i_s)n(X^i_r)1_{X^i_r\in\partial]0,1[} \ge 0$. Therefore

$$\begin{aligned}&[(X^i_r - q(r) - X^i_s + q(s))n(X^i_r)]^-|\mathrm{d}k^i_r|\\&\quad = [(X^i_r - q(r) - X^i_s + q(s))n(X^i_r)]^- 1_{X^i_r\in\partial]0,1[}|\mathrm{d}k^i_r|\\&\quad \le |q(r) - q(s)||\mathrm{d}k^i_r|.\end{aligned}$$

The Lipschitz property of q and Lemma 4.9 give

$$\mathbb{E}\left(\int_s^t |q(r)-q(s)||\mathrm{d}k^i_r|\right)^p \le C|t-s|^p\mathbb{E}\left(\int_0^T |\mathrm{d}k^i_r|\right)^p \le C_p|t-s|^p$$

and we arrive at the second estimate. ∎

Remark 4.11. Only in the above proof do we use Condition 3.1(ii). If μ diverged at the boundaries like $x^{-\alpha}$ for some $\alpha > 1$, then a similar result to Lemma 4.10 should hold, but with $(X^i_r - q(r) - X^i_s + q(s))^\alpha$ in place of $(X^i_r - q(r) - X^i_s + q(s))$. However, such a result would not be enough, since in Lemma 4.12 the power-1 factor $(X^i_r - q(r) - X^i_s + q(s))$ appears and is needed to cancel the interaction term when taking the average. We also expect, for μ diverging like $x^{-\alpha}$ with $\alpha > 1$, that the particles should not even touch the boundaries (as it is without interaction), but we do not focus on this point.

We estimate the Hölder norm of the average of the drift over i.

LEMMA 4.12. *There exists $0 < \alpha \le 1/2$ such that, for every $1 \le p < \infty$, it holds, for some $C_p \ge 0$ independent of s, t, that*

$$\sup_{N,\epsilon}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\int_s^t |\mu^\epsilon(X^{\epsilon,N,i})|\mathrm{d}r\right)^p + \sup_{N,\epsilon}\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\int_s^t |\mathrm{d}k^{\epsilon,N,i}_r|\right)^p \le C_p|t-s|^{\alpha p}.$$

Proof. We start estimating the Hölder norm of $\frac{1}{N}\sum_{i=1}^N |X^i - q|^2$. For this we fix s and we have, for $t > s$,

$$\begin{aligned}(4.5)\qquad &\mathrm{d}|X^i_t - q(t) - X^i_s + q(s)|^2\\&= 2(X^i_t - q(t) - X^i_s + q(s))(-\mu^\epsilon(X^i_t) + \frac{1}{N}\sum_{i=1}^N \mu^\epsilon(X^j_t))\mathrm{d}t\\&\quad + 2\sigma(X^i_t - q(t) - X^i_s + q(s))\left(\mathrm{d}W^i_t - \frac{1}{N}\sum_{j=1}^N \mathrm{d}W^j_t\right) + \sigma^2\left(1-\frac{1}{N}\right)\mathrm{d}t\\&\quad + 2(X^i_t - q(t) - X^i_s + q(s))\left(-\mathrm{d}k^i_t + \frac{1}{N}\sum_{j=1}^N \mathrm{d}k^j_t\right).\end{aligned}$$

Similarly to the argument in Lemma 4.7, averaging over i we get rid of the interaction terms $\frac{1}{N}\sum_{i=1}^{N}\mu^\epsilon(X_t^j)$, $\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}W_t^j$, and $\frac{1}{N}\sum_{j=1}^{N}\mathrm{d}k_t^j$:

$$\begin{aligned}
&\mathrm{d}\frac{1}{N}\sum_{i=1}^{N}|X_t^i - q(t) - X_s^i + q(s)|^2\\
&\quad = -2\frac{1}{N}\sum_{i=1}^{N}(X_t^i - q(t) - X_s^i + q(s))\mu^\epsilon(X_t^i)\mathrm{d}t\\
&\qquad + 2\sigma\frac{1}{N}\sum_{i=1}^{N}(X_t^i - q(t) - X_s^i + q(s))\mathrm{d}W_t^i + \sigma^2\left(1-\frac{1}{N}\right)\mathrm{d}t\\
&\qquad - 2(X_t^i - q(t) - X_s^i + q(s))\mathrm{d}k_t^i.
\end{aligned}$$

We take the p-power and obtain

$$\begin{aligned}
&\left(\frac{1}{N}\sum_{i=1}^{N}|X_t^i - q(t) - X_s^i + q(s)|^2\right)^p\\
&\quad \le C_p\left(\frac{1}{N}\sum_{i=1}^{N}\int_s^t[(X_r^i - q(r) - X_s^i + q(s))\mu^\epsilon(X_r^i)]^-\mathrm{d}r\right)^p\\
&\qquad + C_p\sigma^p\left|\frac{1}{N}\sum_{i=1}^{N}\int_s^t(X_r^i - q(r) - X_s^i + q(s))\mathrm{d}W_r^i\right|^p + C_p\sigma^{2p}|t-s|^p\\
&\qquad + C_p\left(\frac{1}{N}\sum_{i=1}^{N}\int_s^t[(X_r^i - q(r) - X_s^i + q(s))n(X_r^i)]^-|\mathrm{d}k_r^i|\right)^p.
\end{aligned}$$

The first addend of the RHS is controlled via Lemma 4.10 and the Jensen inequality (applied to the average over i),

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\int_s^t[(X_r^i - q(r) - X_s^i + q(s))\mu^\epsilon(X_r^i)]^-\mathrm{d}r\right)^p\\
&\quad \le \sup_{N,\epsilon,i}\mathbb{E}\left(\int_s^t[(X_r^i - q(r) - X_s^i + q(s))\mu^\epsilon(X_r^i)]^-\mathrm{d}r\right)^p \le C_p|t-s|^p,
\end{aligned}$$

and similarly for the last addend. The second addend is controlled via the Burkholder–Davis–Gundy inequality and the Jensen inequality:

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N}\int_s^t(X_r^i - q(r) - X_s^i + q(s))\mathrm{d}W_r^i\right|^p\\
&\quad \le \sup_{N,\epsilon,i}\mathbb{E}\left|\int_s^t(X_r^i - q(r) - X_s^i + q(s))\mathrm{d}W_r^i\right|^p\\
&\quad \le C_p\sup_{N,\epsilon,i}\mathbb{E}\left(\int_s^t|X_r^i - q(r) - X_s^i + q(s)|^2\mathrm{d}r\right)^{p/2}\\
&\quad \le C_p\sup_{N,\epsilon,i}\mathbb{E}\sup_t|X_t^i - q(t)|^p|t-s|^{p/2} \le C_p|t-s|^{p/2}.
\end{aligned}$$

Therefore we have

$$\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t) - X_s^i + q(s)|^2\right)^p \le C_p|t-s|^{p/2}.$$

Now we recall the following elementary inequality (a consequence of Cauchy–Schwarz inequality), for every two sequences of real numbers a_i, b_i:

$$\left|\frac{1}{N}\sum_{i=1}^N (a_i^2 - b_i^2)\right| \le \left(\frac{1}{N}\sum_{i=1}^N |a_i - b_i|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^N |a_i + b_i|^2\right)^{1/2}.$$

Applying this inequality to $a_i = X_t^i - q(t)$, $b_i = X_s^i - q(s)$ and using the Jensen inequality, we get the Hölder bound on $\frac{1}{N}\sum_{i=1}^N |X^i - q|^2$:

(4.6)

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t)|^2 - |X_s^i - q(s)|^2\right|^p \\
&\quad\le \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t) - X_s^i + q(s)|^2\right)^{p/2}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t) + X_s^i - q(s)|^2\right)^{p/2}\right] \\
&\quad\le \left(\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t) - X_s^i + q(s)|^2\right)^p\right)^{1/2}\left(\mathbb{E}\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t) + X_s^i - q(s)|^{2p}\right)^{1/2} \\
&\quad\le C_p|t-s|^{p/4}\sup_{N,\epsilon,i}\left(\mathbb{E}\sup_t |X_t^i - q(t)|^{2p}\right)^{1/2} \le C_p|t-s|^{p/4}.
\end{aligned}$$

On the other hand, averaging (4.4) and using again the cancellation of the interaction terms, we get the equation for $\frac{1}{N}\sum_{i=1}^N |X^i - q|^2$,

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t)|^2 - |X_s^i + q(s)|^2 \\
&\quad = -2\frac{1}{N}\sum_{i=1}^N \int_s^t (X_r^i - q(r))\mu^\epsilon(X_r^i)\mathrm{d}r + 2\frac{1}{N}\sum_{i=1}^N \int_s^t \sigma(X_r^i - q(r))\mathrm{d}W_r^i \\
&\qquad + \sigma^2\left(1 - \frac{1}{N}\right)(t-s) - 2\frac{1}{N}\sum_{i=1}^N \int_s^t (X_r^i - q(r))\mathrm{d}k_r^i,
\end{aligned}$$

and so, by Remark 4.8, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \int_s^t |\mu^\epsilon(X_r^i)|\mathrm{d}r\right)^p + \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \int_s^t |\mathrm{d}k_r^i|\right)^p \\
&\quad\le C_p\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \int_s^t (X_r^i - q(r))\mu^\epsilon(X_r^i)\mathrm{d}r\right)^p + \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \int_s^t (X_r^i - q(r))\mathrm{d}k_r^i\right)^p + C_p(t-s)^p \\
&\quad\le C_p\left|\frac{1}{N}\sum_{i=1}^N |X_t^i - q(t)|^2 - |X_s^i + q(s)|^2\right|^p + C_p\left|\frac{1}{N}\sum_{i=1}^N \int_s^t \sigma(X_r^i - q(r))\mathrm{d}W_r^i\right|^p + C_p(t-s)^p.
\end{aligned}$$

We control the first addend in the RHS by (4.6) and the second addend by the Burkholder–Davis–Gundy inequality (and the Jensen inequality on the average over i):

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\int_s^t|\mu^\epsilon(X_r^i)|\mathrm{d}r\right)^p+\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N\int_s^t|\mathrm{d}k_r^i|\right)^p\\&\le C_p|t-s|^{p/4}+C_p\sup_{N,\epsilon,i}\left(\mathbb{E}\sup_t|X_r^i-q(t)|^p\right)|t-s|^{p/2}+C_p(t-s)^p\le C_p|t-s|^{p/4},\end{aligned}$$

which is the desired estimate with $\alpha=1/4$. ∎

Now we can prove the uniform Hölder bound.

LEMMA 4.13. *With the notation of the previous lemma, for every $1\le p<\infty$, it holds, for some $C_p\ge 0$ independent of s,t, that*

$$\sup_{N,\epsilon,i}\mathbb{E}|X_t^i-X_s^i|^p\le C_p|t-s|^{\alpha p/2}.$$

Proof. By the Jensen inequality, it is enough to prove the estimate for $p\ge 2$. We start again with (4.5) for a fixed i. Taking the $p/2$-power we obtain

$$\begin{aligned}&|X_t^i-q(t)-X_s^i+q(s)|^p\\&\le C_p\left(\int_s^t[(X_r^i-q(r)-X_s^i+q(s))\mu^\epsilon(X_r^i)]^-\mathrm{d}r\right)^{p/2}\\&\quad+C_p\left(\int_s^t|X_r^i-q(r)-X_s^i+q(s)|\frac{1}{N}\sum_{j=1}^N|\mu^\epsilon(X_t^j)|\mathrm{d}r\right)^{p/2}\\&\quad+C_p\sigma^p\left|\int_s^t(X_r^i-q(r)-X_s^i+q(s))\mathrm{d}\left(W_r^i-\frac{1}{N}\sum_{j=1}^N W_r^j\right)\right|^{p/2}+C_p\sigma^{2p}|t-s|^{p/2}\\&\quad+C_p\left(\int_s^t[(X_r^i-q(r)-X_s^i+q(s))n(X_r^i)]^-|\mathrm{d}k_r^i|\right)^{p/2}\\&\quad+C_p\left(\int_s^t|X_r^i-q(r)-X_s^i+q(s)|\frac{1}{N}\sum_{j=1}^N|\mathrm{d}k_r^j|\right)^{p/2}.\end{aligned}$$

The first addend of the RHS is controlled again via Lemma 4.10,

$$\mathbb{E}\left(\int_s^t[(X_r^i-q(r)-X_s^i+q(s))\mu^\epsilon(X_r^i)]^-\mathrm{d}r\right)^{p/2}\le C_p|t-s|^{p/2},$$

and similarly for the fourth addend. The previous Lemma 4.12 allows us to control the second addend,

$$\begin{aligned}&\mathbb{E}\left(\int_s^t|X_r^i-q(r)-X_s^i+q(s)|\frac{1}{N}\sum_{j=1}^N|\mu^\epsilon(X_t^j)|\mathrm{d}r\right)^{p/2}\\&\quad\le C_p\mathbb{E}\left(\int_s^t\frac{1}{N}\sum_{j=1}^N|\mu^\epsilon(X_t^j)|\mathrm{d}r\right)^{p/2}\le C_p|t-s|^{\alpha p/2},\end{aligned}$$

and similarly for the fifth addend. The third addend is controlled via the Burkholder–Davis–Gundy inequality:

$$\mathbb{E}\left(\int_s^t (X_r^i - q(r) - X_s^i + q(s))\mathrm{d}\left(W_r^i - \frac{1}{N}\sum_{j=1}^N W_r^j\right)\right)^{p/2}$$
$$\leq C_p \sup_{N,\epsilon,i} \mathbb{E}\left(\int_s^t |X_r^j - q(r) - X_s^i + q(s)|^2 \mathrm{d}r\right)^{p/4} \leq C_p|t-s|^{p/4}.$$

Putting it all together we get

$$\mathbb{E}|X_t^i - q(t) - X_s^i + q(s)|^p \leq C_p|t-s|^{\alpha p/2}.$$

Using the Lipschitz continuity of q, we obtain the desired bound. ◻

Remark 4.14. We have shown that the Hölder exponent is $\alpha/2 = 1/8$ (since we can take $\alpha = 1/4$). This is a consequence of our argument, but we expect that the optimal Hölder exponent is still $1/2$.

We conclude with a Hölder estimate on the total variation of the drift.

LEMMA 4.15. *For every $1 \leq p < \infty$, it holds for some $C_p \geq 0$ independent of s, t that*

$$\sup_{N,\epsilon,i} \mathbb{E}\left(\int_s^t \left|\mu^\epsilon\left(X_r^{\epsilon,N,i}\right)\right| \mathrm{d}r\right)^p + \sup_{N,\epsilon,i} \mathbb{E}\left(\int_s^t \left|\mathrm{d}k_r^{\epsilon,N,i}\right|\right)^p \leq C|t-s|^{\alpha p/2}.$$

Proof. Equation (4.4) for $|X_t - q(t)|^2$ implies

$$\int_s^t 2(X^i - q(r))\mu^\epsilon(X^i)\mathrm{d}r + \int_s^t 2(X^i - q(r))\mathrm{d}k_r^i$$
$$\leq |X_s^i - q(s)|^2 - |X_t^i - q(t)|^2 + C\int_s^t \frac{1}{N}\sum_{j=1}^N |\mu^\epsilon(X^j)|\mathrm{d}r$$
$$+ \left|\int_s^t 2\sigma(X^i - q(t))\left(\mathrm{d}W^i - \frac{1}{N}\sum_{j=1}^N \mathrm{d}W^j\right)\right| + \sigma^2(t-s) + C\int_s^t \frac{1}{N}\sum_{i=1}^N |\mathrm{d}k_r^j|,$$

and so, by Remark 4.8,

$$\int_s^t |\mu^\epsilon(X^i)|\mathrm{d}r + \int_s^t |\mathrm{d}k_r^{\epsilon,N,i}|$$
$$\leq C|X_s^i - X_t^i - q(s) + q(t)||X_s^i + X_t^i - q(s) - q(t)| + C\int_s^t \frac{1}{N}\sum_{j=1}^N |\mu^\epsilon(X^j)|\mathrm{d}r$$
$$+ C\left|\int_s^t \sigma(X^i - q(t))\left(\mathrm{d}W^i - \frac{1}{N}\sum_{j=1}^N \mathrm{d}W^j\right)\right| + C\sigma^2(t-s) + C\int_s^t \frac{1}{N}\sum_{j=1}^N |\mathrm{d}k_r^j|.$$

By the Burkholder–Davis–Gundy inequality, we get

$$\mathbb{E}\left(\int_s^t |\mu^\epsilon(X^i)|\mathrm{d}r\right)^p + \mathbb{E}\left(\int_s^t |\mathrm{d}k_r^{\epsilon,N,i}|\right)^p$$
$$\leq C\mathbb{E}|X_s^i - X_t^i - q(s) + q(t)|^p + C\mathbb{E}\left(\int_s^t \frac{1}{N}\sum_{j=1}^N |\mu^\epsilon(X^j)|\mathrm{d}r\right)^p$$
$$+ C(t-s)^p + C\mathbb{E}\left(\int_s^t \frac{1}{N}\sum_{j=1}^N |\mathrm{d}k_r^j|\right)^p.$$

Lemmas 4.12 and 4.13 allow us to conclude the desired bound. ∎

4.4. Convergence of the particle system. In this subsection we show the convergence of the regularized particle system (3.4) to the McKean–Vlasov SDE (3.5), by a compactness argument; as a consequence, we get the existence of a solution to (3.5). We are given a probability space $(\Omega, \mathcal{A}, P)$ and independent Brownian motions W^i, $i \geq 1$, on a (right-continuous complete) filtration $(\mathcal{F}_t)_t$. For each N, we are given an $(X_0^{1,N}, \dots X_0^{N,N})$ $\mathcal{F}_0$-measurable random variable and we let $(X^{i,N,\epsilon}, k^{i,N,\epsilon})$ be the corresponding solution to the regularized N-particle system (3.4). Throughout the section, we assume Conditions 3.1, 3.2, and 3.3.

In the following we use the notation $C_t = C([0,T];\mathbb{R}^d)$, $C_{t,[0,1]} = C([0,T];[0,1])$; we use also $W_t^{\beta,p} = W^{\beta,p}([0,T])$ for the fractional Sobolev space of order $0 < \beta < 1$ and exponent $1 \leq p < \infty$, with norm

$$\|f\|_{W_t^{\beta,p}}^p = \int_0^T |f(t)|^p\mathrm{d}t + \int_0^T\int_0^T \frac{|f(t)-f(s)|^p}{|t-s|^{1+\beta p}}\mathrm{d}s\mathrm{d}t.$$

We consider the Polish space $E = C_t \times C_{t,[0,1]} \times C_t$, endowed with its Borel σ-algebra. We denote a generic element of E as $\gamma = (\gamma^1, \gamma^2, \gamma^3)$ or (for reasons that will be clear later) (W, X, Z); with a little abuse of notation, we use

$$W, X, Z$$

also to denote the canonical projections on E. We consider also the space $\mathcal{P}(E)$ of probability measures on E with the topology of weak convergence of probability measure (also endowed with its Borel σ-algebra). The space $\mathcal{P}(E)$ is a Polish space as well [AGS08, Remark 7.1.7]. For a measure ν and a function g on E, we use the notation $\nu(g) = \int_E g\mathrm{d}\mu$ (when the integral makes sense).

We take the random empirical measures on E given by

$$L^{N,\epsilon} = \frac{1}{N}\sum_{i=1}^N \delta_{\left(W^i(\omega), X^{i,N,\epsilon}(\omega), -\int_0^\cdot \mu^\epsilon(X_r^{i,N,\epsilon})\mathrm{d}r - k^{i,N,\epsilon}\right)},$$

which are random variables on $\mathcal{P}(E)$. Note that, for ω in Ω, for any Borel bounded or nonnegative function $g: E \to \mathbb{R}$,

$$\mathbb{E}^{L^{N,\epsilon}(\omega)}[g(W,X,Z)] = \frac{1}{N}\sum_{i=1}^N g\left(W^i(\omega), X^{i,N,\epsilon}(\omega), -\int_0^\cdot \mu^\epsilon\left(X_r^{i,N,\epsilon}\right)\mathrm{d}r - k^{i,N,\epsilon}\right),$$

where $\mathbb{E}^{L^{N,\epsilon}}$ is the expectation under $L^{N,\epsilon}$. By (3.4) and the definition of $L^{\epsilon,N}(\omega)$, for P-a.e. ω, under the measure $L^{N,\epsilon}(\omega)$ it holds on E that for every t,

$$
\begin{aligned}
&X_t = X_0 + q(t) - q(0) + \sigma W_t + Z_t - \sigma\mathbb{E}^{L^{N,\epsilon}}[W_t] - \mathbb{E}^{L^{N,\epsilon}}[Z_t],\\
&Z_t = -\int_0^t \mu^\epsilon(X_r)\mathrm{d}r - k_t,\\
&\mathrm{d}|k|_t = 1_{X_t\in\{0,1\}}\mathrm{d}|k|_t, \quad \mathrm{d}k_t = n(X_t)\mathrm{d}|k|_t.
\end{aligned}
\tag{4.7}
$$

PROPOSITION 4.16. *Assume Conditions* 3.1*,* 3.2*, and* 3.3 *(actually, Condition* 3.1(iii) *is not needed). Then the family* $(Law(L^{N,\epsilon}))_{N,\epsilon}$ *(probability measures on* $\mathcal{P}(E)$*) is tight.*

Remark 4.17. In view of the proof, we recall the following standard/known facts:

- To prove that a family of probability measures $(P^n)_n$ on a metric space χ is tight, it is enough to find a nonnegative function F on χ such that F is coercive (that is, with compact sublevel sets) and $\int_\chi F\mathrm{d}P^n$ is bounded uniformly in n.
- When $\chi = \mathcal{P}(E)$ for E as above (and more generally for every Polish space E), endowed with the topology of weak convergence, we can take $F(\nu) = \int_E g\mathrm{d}\nu$ as a nonnegative coercive function on $\mathcal{P}(E)$, provided that $g : E \to \mathbb{R}$ is a nonnegative coercive function on E. Indeed, every sublevel set $\{F \le C\}$ is compact: for any sequence $(\nu^n)_n$ of measures on E, if all ν^n belong to $\{F \le C\}$, then, by the previous point applied to g, $(\nu^n)_n$ is tight, hence there exists a subsequence which is weakly convergent to some measure ν on E, and μ also belongs to $\{F \le C\}$ by the Fatou lemma.
- By Sobolev embedding, there exists $C > 0$ such that, for every γ in $E = C_t \times C_{t,[0,1]} \times C_t$,

  $$\|\gamma\|_{C_t^\alpha} \le C\|\gamma\|_{W_t^{\beta,p}}$$

  for $\alpha = \beta - 1/p$, provided that $\beta - 1/p > 0$. By the Ascoli–Arzelà theorem, the norm $\|\cdot\|_{C_t^\alpha}$ is coercive on E for $\alpha > 0$. Therefore, to show that a certain family of probability measures P^n on $\mathcal{P}(E)$ is tight, it is enough to show that

  $$\sup_n \mathbb{E}^{P^n} \int_E \|\gamma\|_{W_t^{\beta,p}}\mu(\mathrm{d}\gamma) < \infty$$

  for some $\beta > 0$, $p \ge 1$ with $\beta - 1/p > 0$ (here $\mathbb{E}^{P^n}$ denotes the expectation under P^n).

Proof. By the previous Remark 4.17, it is enough to verify that, for some $\beta > 0$, $p \ge 1$ with $\beta > 1/p$, for $h = 1, 2, 3$,

$$\sup_n \mathbb{E}\int_E \|\gamma^h\|_{W_t^{N,\epsilon}} L^{\epsilon,N}(\mathrm{d}\gamma) < \infty.$$

For $h = 1$, that is, the Brownian motion component, we have

$$\mathbb{E}\int_E \|\gamma^1\|_{W_t^{\beta,p}} L^{N,\epsilon}(\mathrm{d}\gamma) = \mathbb{E}\frac{1}{N}\sum_{i=1}^N \|W^i\|_{W_t^{\beta,p}} = \mathbb{E}\|W^1\|_{W_t^{\beta,p}} < \infty$$

for any $\beta < 1/2$ and $p \ge 1$.

For $h = 2$, that is, the solution component, by Lemma 4.13 we get, for some $\beta > 0$, for every $p \geq 1$ and $0 < \delta < \beta$, for every $i = 1, \dots N$,

$$\begin{aligned}
&\mathbb{E}[\|X^{i,N,\epsilon}\|^p_{W_t^{\beta-\delta,p}}] \\
&= \int_0^T \mathbb{E}|X_t^{i,N,\epsilon}|^p \mathrm{d}t + \int_0^T \int_0^T \frac{\mathbb{E}|X_t^{i,N,\epsilon} - X_s^{i,N,\epsilon}|^p}{|t-s|^{1+(\beta-\delta)p}} \mathrm{d}s\mathrm{d}t \\
&\leq T + \int_0^T \int_0^T |t-s|^{-(1-\delta p)} \mathrm{d}s\mathrm{d}t \sup_{s,t} \frac{\mathbb{E}|X_t^{i,N,\epsilon} - X_s^{i,N,\epsilon}|^p}{|t-s|^{\beta p}} \leq C
\end{aligned}$$

for some $C > 0$ independent of N and ϵ. It follows that

$$\mathbb{E}\|\gamma^2\|_{W_t^{\beta-\delta\delta,p}} = \mathbb{E}\frac{1}{N}\sum_{i=1}^N \|X^{i,N,\epsilon}\|_{W_t^{\beta-\delta,p}} \leq C.$$

A similar argument, using Lemma 4.15 in place of Lemma 4.13, works for $h = 3$. The proof is complete. ∎

From now on, we assume that $\frac{1}{N}\sum_{i=1}^N \delta_{X_0^{i,N}}$ converges in law to some probability measure $\mathrm{Law}(X_0)$.

In the following, we fix a limit point Q of $\mathrm{Law}(L^{N,\epsilon})$ and a $\mathcal{P}(E)$-valued random variable L with law Q. With a little abuse of notation, we do not relabel the subsequence of $\mathrm{Law}(L^{N,\epsilon})$ converging to Q, and we assume that L is also defined on the same probability space $(\Omega, \mathcal{A}, P)$ (this is only for notational simplicity). We call $Q^{N,\epsilon,e}$, Q^e the (deterministic) probability measures on E obtained averaging, resp., $L^{N,\epsilon}$, L, namely, for every Borel bounded or nonnegative function g on E,

$$\mathbb{E}^{Q^{\epsilon,N,e}}[g] = \mathbb{E}[L^{\epsilon,N}(g)], \quad \mathbb{E}^{Q^e}[g] = \mathbb{E}[L(g)],$$

where $\mathbb{E}^Q$ denotes the expectation with respect to Q.

Remark 4.18. We recall some useful facts of convergence in the law of random probability measures.

- Let $H : E \to \tilde{E}$ a continuous map with values in some Polish space $\tilde{E}$; then the $\mathcal{P}(\tilde{E})$-valued random variables $H_\# L^{N,\epsilon}$ converge in law to $H_\# L$. This follows from the continuity of the map $\nu \mapsto H_\#\nu$, which in turn follows from the continuity of H.
- Let g be in $C_b(E)$; then the real-valued random variables $L^{N,\epsilon}(g)$ converge in law to $L(g)$. Similarly to the previous point, this follows from the continuity of the map $\nu \mapsto \nu(g)$.
- The probability measures $Q^{N,\epsilon,e}$ converge weakly to Q^e: indeed, for every g in $C_b(E)$, $\mathbb{E}[L^{N,\epsilon}(g)]$ converge to $\mathbb{E}[L(g)]$.
- For any Borel set B of E, it holds that $L^{N,\epsilon}(B) = 1$ P-a.s. if and only if $Q^{N,\epsilon,e}(B) = 1$, and similarly for L and Q^ϵ.

Now we show that, for a.e. L, W (the first component in E) is a Brownian motion under L and X is a generalized solution to the McKean–Vlasov equation, with the right initial condition, under L. Roughly speaking, we would like to pass to the limit (as $N \to \infty$ and $\epsilon \to 0$) in (4.7) and get the McKean–Vlasov SDE (3.6). The proof is in two parts. In the first part, we prove that the expectation of X_t under L is $q(t)$, which implies the first line in (3.6); we also prove that the law of the initial condition X_0 and W is $\mathrm{Law}(X_0) \otimes$Wiener measure. In the second part we idenfity the reflection term and prove its properties, getting the equalities for $\bar{Z}$ and $\bar{k}$ in (3.6).

LEMMA 4.19. *It holds P-a.s. that*

- *for every t,*

$$\mathbb{E}^L[X_t] = q(t);$$

- *under the measure L, the random path*

$$t \mapsto X_t - X_0 - \sigma W_t - Z_t$$

is actually L-a.s. deterministic; that is, the law of this path under L is a Dirac delta.

Proof. For the first point, we start fixing t. The function on E defined by $(W,X,Z) \mapsto X_t$ is continuous and bounded (as X takes values in $[0,1]$). Therefore, by Remark 4.18, the random variables $\mathbb{E}^{L^{N,\epsilon}}[X_t]$ converge in law to $\mathbb{E}^L[X_t]$. On the other hand, (3.4) gives, for every (N,ϵ), P-a.s.,

$$\mathbb{E}^{L^{N,\epsilon}}[X_t] = \frac{1}{N}\sum_{i=1}^N X^{i,N,\epsilon} = q(t).$$

Hence the law of $\mathbb{E}^L[X_t]$ is $\delta_{q(t)}$, that is, $\mathbb{E}^L[X_t] = q(t)$ on a P-full measure set Ω_t, which may depend on t. To make the exceptional set independent on t, we note that, by the dominated convergence theorem, $t \mapsto \mathbb{E}^L[X_t]$ is continuous for every ω and that q is also continuous by assumption, hence we have the equality for every t in the full-measure set $\Omega' = \cap_{s\in\mathbb{Q}\cap[0,T]}\Omega_s$. The proof of the first point is complete.

For the second point, we have to prove that $(X-X_0-\sigma W-Z)_\# L$ is a Dirac delta P-a.s. By Remark 4.18, the $\mathcal{P}(C_t)$-valued random variables $(X-X_0-\sigma W-Z)_\# L^{N,\epsilon}$ converge in law to $(X-X_0-\sigma W-Z)_\# L$. On the other hand, (3.4) gives that, for every (N,ϵ), P-a.s., for every $i=1,\dots N$,

$$\begin{aligned}
&(X - X_0 - \sigma W - Z)_\# \delta_{\left(W^i, X^{i,N,\epsilon}, -\int_0^\cdot \mu^\epsilon(X_r^{i,N,\epsilon})\mathrm{d}r - k^{i,N,\epsilon}\right)} \\
&\quad = \delta_{X^{i,N,\epsilon} - X_0^{i,N,\epsilon} - \sigma W_t^i + \int_0^\cdot \mu^\epsilon\left(X_r^{i,N,\epsilon}\right)\mathrm{d}r + k^{i,N,\epsilon}} \\
&\quad = \delta_{q(t)-q(0)+\frac{1}{N}\sum_{j=1}^N\left[\int_0^\cdot \mu^\epsilon\left(X_r^{j,N,\epsilon}\right)\mathrm{d}r - \sigma W_t^j + k^{j,N,\epsilon}\right]} =: \delta_{\gamma^{N,\epsilon}};
\end{aligned}$$

note that $\gamma^{N,\epsilon}$ is independent of i. Averaging over i, we get

$$(X - X_0 - \sigma W - Z)_\# L^{N,\epsilon} = \delta_{\gamma^{N,\epsilon}},$$

in particular $(X-X_0-\sigma W-Z)_\# L^{N,\epsilon}$ is concentrated on the subset $\{\delta_\gamma \mid \gamma \in C_t\}$ of $\mathcal{P}(C_t)$. We claim that $\{\delta_\gamma \mid \gamma\in C_t\}$ is a closed set in $\mathcal{P}(C_t)$. Hence, since $(X-X_0-\sigma W-Z)_\# L^{N,\epsilon}$ converges in law to $(X-X_0-W-Z)_\# L$, also $(X-X_0-\sigma W-Z)_\# L$ is concentrated on $\{\delta_\gamma \mid \gamma\in C_t\}$, that is, the law of $X-X_0-\sigma W-Z$ under L is a Dirac delta.

It remains to prove the above claim. If δ_{γ^n} converge to a measure ν, then, by tightness of δ_{γ^n}, there exists a compact set K in C_t such that $\delta_{\gamma^n}(K) > 1/2$ and so γ^n belong to K, for every n. Therefore there exists a subsequence γ^{n_k} converging to some element γ in K, hence $\delta_{\gamma^{n_k}}$ converge to δ_γ and so $\nu = \delta_\gamma$ belongs to $\{\delta_\gamma \mid \gamma \in C_t\}$, which is then closed. The proof of the second point is complete. ∎

LEMMA 4.20. *It holds P-a.s. that under L, the $C_t \times \mathbb{R}$-valued random variable (W, X_0) has law $P^W \otimes Law(X_0)$, where P^W is the Wiener measure.*

Proof. The map from E to $C_t \times \mathbb{R}$ defined by $(W, X, Z) \mapsto (W, X_0)$ is continuous. Therefore, by Remark 4.18, the random empirical measures

$$(W, X_0)_\# L^{N,\epsilon} = \frac{1}{N} \sum_{i=1}^N \delta_{\left(W^i, X_0^{i,N}\right)}$$

converge in law to $(W, X_0)_\# L$. On the other hand, the above random measures converge in law to $P^W \otimes \mathrm{Law}(X_0)$ (see, e.g., [CDFM20, Lemma 29]). Hence the law of $\mathrm{Law}^L(W, X_0)$ is $\delta_{P^W \otimes \mathrm{Law}(X_0)}$, that is, $\mathrm{Law}^L(W, X_0) = P^W \otimes \mathrm{Law}(X_0)$ P-a.s. The proof is complete. ∎

Next we define the process k by

$$k_t = k(X, Z)_t = -\int_0^t 1_{X_r \notin]0,1[} \mathrm{d}Z_r \tag{4.8}$$

if Z is a BV path on $[0, T]$, $k_t = 0$ otherwise. We call $|k|$ the total variation process associated with k.

LEMMA 4.21. *The following hold P-a.s.:*

- *The processes Z, $\int_0^\cdot \mu(X_r)\mathrm{d}r$, k have BV trajectories L-a.e. and their BV norms are p-integrable with respect to L, for any $1 \le p < \infty$.*
- *It holds L-a.e. that for every $t \ge 0$,*

$$Z_t + k_t = \int_0^t 1_{X_r \in]0,1[} \mathrm{d}Z_r = -\int_0^t \mu(X_r) 1_{X_r \in]0,1[} \mathrm{d}r. \tag{4.9}$$

- *It holds L-a.e. that the process k satisfies the condition*

$$\mathrm{d}|k|_t = 1_{X_t \in \{0,1\}} \mathrm{d}|k|_t, \quad \mathrm{d}k_t = n(X_t)\mathrm{d}|k|_t. \tag{4.10}$$

Proof. For all statements but the p-integrability of the BV norms, by Remark 4.18, it is enough to prove these statements Q^e-a.e. (recall Q^e is the average of L) instead of L-a.e. (provided we work with Borel sets/properties, as the proof will do); it is also enough to prove p-integrability of the BV norms of Z, $\int_0^\cdot \mu(X_r)\mathrm{d}r$, k with respect to Q^e. Again by Remark 4.18, the measures $Q^{N,\epsilon,e}$ converge in law to Q^e, hence we can work with $Q^{N,\epsilon,e}$ and Q^e only.

BV property of Z and k. By Lemma 4.9, we have

$$\mathbb{E}^{Q^{N,\epsilon,e}} \|Z\|_{BV}^p \le \mathbb{E} \frac{1}{N} \sum_{i=1}^N \left(\int_0^T \left| \mu^\epsilon \left(X_r^{i,N,\epsilon} \right) \right| \mathrm{d}r + \left| k^{i,N,\epsilon} \right|_T \right)^p \le C \tag{4.11}$$

for some constant C independent of ϵ and N. Now the BV norm is lower semicontinuous in C_t, since it can be written as

$$\|\gamma\|_{BV} = \sup_\pi \sum_{[t_i, t_{i+1}[\in \pi} |\gamma(t_{i+1}) - \gamma(t_i)|,$$

the sup being over all partitions π of $[0, T]$. Therefore it holds that

$$\mathbb{E}^{Q^e} \|Z\|_{BV}^p \le C,$$

in particular Z, and so k have BV paths Q^e-a.s.

Support property of $|k|$. By definition, k is concentrated on $\{t \in [0,T] \mid X_t \in \{0,1\}\}$, which is a closed set, hence also its total variation process $|k|$ is concentrated on this set and we conclude that

$$\mathrm{d}|k|_t = 1_{X_t\in\{0,1\}}\mathrm{d}|k|_t. \tag{4.12}$$

BV property of $\int_0^\cdot \mu(X_r)\mathrm{d}r$. Since $\mu^\epsilon \geq \mu^\delta$ for $\epsilon < \delta$, by the monotone convergence theorem we have

$$\begin{aligned}
\mathbb{E}^{Q^e}\left(\int_0^T |\mu(X_r)|\mathrm{d}r\right)^p &= \sup_\delta \mathbb{E}^{Q^e}\left(\int_0^T \left|\mu^\delta(X_r)\right|\mathrm{d}r\right)^p\\
&= \sup_\delta \lim_{N,\epsilon}\left(\mathbb{E}^{Q^{N,\epsilon,e}}\int_0^T \left|\mu^\delta(X_r)\right|\mathrm{d}r\right)^p\\
&= \sup_\delta \lim_{N,\epsilon}\mathbb{E}\frac{1}{N}\sum_{i=1}^N\left(\int_0^T \left|\mu^\delta\left(X_r^{N,\epsilon,i}\right)\right|\mathrm{d}r\right)^p\\
&\leq \liminf_{N,\epsilon}\mathbb{E}\frac{1}{N}\sum_{i=1}^N\left(\int_0^T \left|\mu^\epsilon\left(X_r^{N,\epsilon,i}\right)\right|\mathrm{d}r\right)^p < \infty;
\end{aligned}$$

in particular also $\int_0^\cdot \mu(X_r)\mathrm{d}r$ has BV trajectories, with Q^e-integrable BV norm.

Representation formulae for $Z+k$ *and* k. We take $a>0$, $\delta>0$, $\varphi:[0,1]\to\mathbb{R}$ C^1 function with support in $[\delta,1-\delta]$, $\tilde n;[0,1]\to\mathbb{R}$ a continuous extension of the outer normal n with support on $]\delta,1-\delta[^c$ and with $\tilde n\geq 0$ on $[1-\delta,1]$ and $\tilde n\leq 0$ on $[0,\delta]$, $g,h:[0,T]\to\mathbb{R}$ continuous with g nonnegative. We consider the set

$$\begin{aligned}
A &= A_{a,\varphi,\tilde n,h,g}\\
&= \Bigg\{(W,X,Z)\in E \mid \|Z\|_{BV}\leq a,\ \int_0^T h(r)\varphi(X_r)\mathrm{d}Z_r\\
&\qquad = -\int_0^T h(r)\varphi(X_r)\mu(X_r)\mathrm{d}r,\ \int_0^T g(r)\tilde n(X_r)\mathrm{d}Z_r \leq 0\Bigg\}.
\end{aligned}$$

LEMMA 4.22. *The set* A *is closed in* E.

Proof. Let (W^n,X^n,Z^n) be a sequence in A converging to (W,X,Z) uniformly. Since the BV norm of Z^n is bounded by a for every n, up to taking a subsequence we can assume that dZ^n converges weakly-* to a measure ν with total variation $\|\nu\|_{TV}\leq a$. Passing to the limit in the chain rule for Z, we find that, for every ψ in $C^\infty([0,T])$,

$$\int_0^T \psi\mathrm{d}\nu = \psi(T)Z_T - \psi(0)Z_0 - \int_0^T \psi' Z\mathrm{d}r.$$

Hence ν is the distributional derivative of Z, which therefore satisfies $\|Z\|_{BV}\leq a$.

Concerning the stability of the conditions involving μ and $\tilde n$, note that $h(r)\varphi(X^n_r)\to h(r)\varphi(X_r)$ uniformly and also $h(r)\varphi(X^n_r)\mu(X^n_r)\to h(r)\varphi(X_r)\mu(X_r)$ uniformly (since μ is C^1 on $[\delta,1-\delta]$). This fact and the weak-* convergence of Z^n imply that

$$\int_{[0,T]} h(r)\varphi(X_r)\mathrm{d}Z_r = \lim_n \int_0^T h(r)\varphi(X_r^n)\mathrm{d}Z_r^n$$
$$= -\lim_n \int_0^T h(r)\varphi(X_r^n)\mu(X_r^n)\mathrm{d}r = -\int_0^T h(r)\varphi(X_r)\mu(X_r)\mathrm{d}r.$$

Reasoning similarly for $\tilde{n}$, we find

$$\int_0^T g(r)\tilde{n}(X_r)\mathrm{d}Z_r = \lim_n \int_0^T g(r)\tilde{n}(X_r^n)\mathrm{d}Z_r^n \leq 0.$$

This proves that (W, X, Z) is in A. Hence A is closed. ∎

Now the equation for $X^{\epsilon,N}$ and Condition 3.1(iii) imply, for $\delta < \rho$, that for $\epsilon < \delta$, under $Q^{\epsilon,N,e}$ it holds a.s.

$$\int_0^T h(r)\varphi(X_r)\mathrm{d}Z_r = -\int_0^T h(r)\varphi(X_r)\mu(X_r)\mathrm{d}r,$$
$$\int_0^T g(r)\tilde{n}(X_r)\mathrm{d}Z_r = -\int_0^T g(r)\tilde{n}(X_r)(\mu^\epsilon(X_r)\mathrm{d}r + \mathrm{d}k_r) \leq 0.$$

Moreover the uniform bound (4.11) implies

$$Q^{\epsilon,N,e}\{\|Z\|_{BV} > a\} \leq \frac{1}{a}\mathbb{E}^{Q^{\epsilon,N,e}}\|Z\|_{BV} \leq C/a.$$

Therefore, for $\delta < \rho$, for any a, $Q^{\epsilon,N,e}(A) \geq 1 - C/a$. Since A is closed, we conclude that $Q^e(A) \geq 1 - C/a$. Hence Q^e is concentrated on the set

$$B_{\varphi,\tilde{n},h,g} = \left\{(W, X, Z) \in \Omega \mid \|Z\|_{BV} < \infty,\ \int_0^T h(r)\varphi(X_r)\mathrm{d}Z_r = -\int_0^T h(r)\varphi(X_r)\mu(X_r)\mathrm{d}r,\ \int_0^T g(r)\tilde{n}(X_r)\mathrm{d}Z_r \leq 0\right\}$$

for every φ, $\tilde{n}$, h, g as above. Now we take $\varphi = \varphi^m$ tending pointwise to $1_{]0,1[}$ and uniformly bounded in m; $\tilde{n} = \tilde{n}^m$ tending pointwise to $n(x)1_{\{0,1\}}$ and uniformly bounded in m; and h in acountable dense set D in C_t and g in D^+ countable dense set D^+ in $\{g \in C_t \mid g \geq 0\}$. Therefore we have

(4.13)
$$Q^e \text{ is concentrated on } \tilde{B} = \cap_{m,h\in D,g\in D^+} B_{\varphi^m,\tilde{n}^m,h,g} \cap \left\{(W, X, Z) \mid \int_0^\cdot \mu(X)\mathrm{d}r \in BV\right\}.$$

LEMMA 4.23. *For every (W, X, Z) in $\tilde{B}$, it holds that*

$$\int_0^t 1_{X_r\in]0,1[}\mathrm{d}Z_r = -\int_0^t \mu(X_r)1_{X_r\in]0,1[}\mathrm{d}r \quad \forall t, \tag{4.14}$$

$$\mathrm{d}k_t = n(X_t)\mathrm{d}|k|_t. \tag{4.15}$$

Proof. For every (W, X, Z), for every fixed h in D and g in D^+, the BV property of Z (and so of k) and of $\int_0^\cdot \mu(X_r)\mathrm{d}r$ implies, via the dominated convergence theorem,

$$\int_0^T h(r)\varphi^m(X_r)\mathrm{d}Z_r \to \int_0^T h(r)1_{X\in]0,1[}\mathrm{d}Z_r, \quad \int_0^T h(r)\varphi^m(X_r)\mu(X_r)\mathrm{d}r \to \int_0^T h(r)1_{X\in]0,1[}\mu(X_r)\mathrm{d}r,$$

$$\int_0^T g(r)\tilde{n}^m(X_r)\mathrm{d}Z_r \to -\int_0^T g(r)n(X_r)\mathrm{d}k_r.$$

Therefore, if (W, X, Z) is in $\tilde{B}$, passing to the limit in m in the definition of $B_{\varphi,\tilde{n},h,g}$ we get

$$\int_0^T h(r)1_{X\in]0,1[}\mathrm{d}Z_r = -\int_0^T h(r)\mu(X_r)1_{X\in]0,1[}\mathrm{d}r - \int_0^T g(r)n(X_r)\mathrm{d}k_r \leq 0$$

for all h in D, g in D^+. By the density of D and D^+ we obtain (4.14) and that $n(X_r)dk_r \geq 0$, which together with (4.12) implies (4.15). ∎

Thanks to (4.13) and Lemma 4.23, we conclude that, for Q^e-a.e. (W, X, Z), the representation formulae (4.14) and (4.15) hold. Therefore (4.9) and (4.10) hold. The proof of Lemma 4.21 is complete. ∎

Remark 4.24. Only in the above proof do we use Condition 3.1(iii). If $\sigma \neq 0$, we expect that, by a suitable version of the Girsanov theorem on domains, the time spent by X on the boundary has zero Lebesgue measure, Q-a.s. Essentially this should allow us to remove or relax Condition 3.1(iii).

We are ready to prove the following.

PROPOSITION 4.25. *It holds P-a.e. that under L, (X, k) is a generalized solution to the McKean–Vlasov problem* (3.5) *starting from X_0, with initial distribution $Law(X_0)$ (more precisely, $(E, \mathcal{B}(E), W, X, k, L)$ is a weak generalized solution with initial distribution $Law(X_0)$).*

Proof. By Lemma 4.20, P-a.s., W is a Brownian motion under L and X_0 is independent of W. As a consequence of Lemma 4.19 it holds that, P-a.s., under L, for every t,

$$X_t = X_0 + Z_t + \sigma W_t + q(t) - q(0) - \mathbb{E}^L Z_t, \tag{4.16}$$

where we have used $\mathbb{E}^L W_t = 0$. By Lemma 4.21, it holds that, P-a.s., under L, $\int_0^\cdot \mu(X_r)1_{X_r\in]0,1[}\mathrm{d}r$ and k are in BV with integrable BV norms and, for every t,

$$Z_t = -\int_0^t \mu(X_r)1_{X_r\in]0,1[}\mathrm{d}r - k_t = -\int_0^t \mu(X_r)\mathrm{d}r - k_t, \tag{4.17}$$

where k_t satisfies (4.10) and where we have used that $\mu(0) = \mu(1) = 0$. Therefore (X, k) satisfies (3.6) and so it is a generalized solution. ∎

We deduce, via Yamada–Watanabe, the existence of a strong solution to (3.5), that is the existence part of Theorem 3.9, as well as uniqueness in law.

COROLLARY 4.26. *It holds that P-a.s., under L, (X,k) is a strong solution to the SDE* (3.5) *and the law of X under L coincides with the unique law $Law(\bar{X})$ of any solution to* (3.5) *starting from $Law(\bar{X}_0) = Law(X_0)$.*

Proof. We have P-a.s. that the couple (X,k) is a weak generalized solution under L, hence it is a strong solution, via Proposition 4.6. Proposition 4.5 gives uniqueness in law for the X component. ∎

Finally we arrive at the convergence result, that is, Theorem 3.10.

COROLLARY 4.27. *The family $(\frac{1}{N}\sum_{i=1}^N \delta_{X^{\epsilon,N,i}} = X_\# L^{\epsilon,N})_{\epsilon,N}$ of random probability measures on $C([0,T];[0,1])$ converges in probability, as $\epsilon \to 0$ and $N \to \infty$, to the law of the McKean–Vlasov solution $\bar{X}$ (starting from $Law(X_0)$).*

Proof. Since the limit $\mathrm{Law}(\bar{X})$ is deterministic (and $\mathcal{P}(E)$ is a metric space), it is enough to prove convergence in law. Since $\mathcal{P}(\mathcal{P}(E))$ is a metric space and the family $(\mathrm{Law}(X_\# L^{\epsilon,N})_{\epsilon,N})_{\epsilon,N}$ is relatively compact (that is, tight), it is enough to prove that every limit point of $(\mathrm{Law}(X_\# L^{\epsilon,N})_{\epsilon,N})_{\epsilon,N}$ is actually $\delta_{\mathrm{Law}(\bar{X})}$. This is an immediate consequence of Corollary 4.26. The proof is complete. ∎

4.5. Pathwise analysis. This subsection is dedicated to the proof of Proposition 3.12; we assume in this subsection the conditions of Proposition 3.12. We use a pathwise approach developed, e.g., in [CDFM20]; we explain first briefly the core idea behind it. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probablity space and $W : \Omega \to C([0,T],\mathbb{R})$ a random variable on this space. Note that at this point we do not impose that W is a Brownian motion. Consider the SDE

$$\begin{aligned} &\mathrm{d}X_t = [-\mu(X_t) + \mathbb{E}[\mu(X_t)]]\mathrm{d}t + \dot{q}_t\mathrm{d}t + \sigma\mathrm{d}W_t - \mathrm{d}k_t - \sigma\mathbb{E}[\mathrm{d}W_t] + \mathbb{E}[\mathrm{d}k_t],\\ &X \in C([0,T];[0,1]),\ k \in C([0,T];\mathbb{R}),\\ &\mathrm{d}|k| = 1_{X_t\in\{0,1\}}\mathrm{d}|k|,\quad \mathrm{d}k = n(X_t)\mathrm{d}|k|. \end{aligned} \tag{4.18}$$

If we endow the probability space with a (right-continuous complete) filtration $(\mathcal{F}_t)_{t\geq 0}$ and assume that W is a Brownian motion with respect to this filtration, clearly (4.18) is exactly the McKean–Vlasov equation (3.5).

On the other hand, let $(X^{(N)}, k^{(N)})$ be the solution of the interacting particle system (3.4). Let $\omega \in \Omega$ be fixed. On a suitable discrete prabability space endowed with the point counting measure the process $(X^{(N)}, k^{(N)})(\omega) = (X^{i,N}(\omega), k^{i,N}(\omega))_{i=1,\dots,N}$ is a random variable in the variable i and as such a solution to (4.18). The mean with respect to the point counting measure is exactly the empirical average.

This is the main idea behind the proof of the Lemma 4.29. First we recall the definition of Wasserstein distance.

DEFINITION 4.28. *Let (E,d) be a Polish space. Let $\mathcal{P}_2(E)$ be the space of probability measures on E with finite second moment. The 2-Wasserstein distance on $\mathcal{P}(E)$ is defined as*

$$\mathcal{W}_{2,E}(\mu,\nu) := \inf\left\{\left(\int_{E\times E} d(x,y)^2 m(dx,dy)\right)^{\frac{1}{2}} \mid m \text{ coupling of } \mu,\nu\right\},\quad \mu,\nu\in\mathcal{P}_2(E).$$

From now on, we work under the assumptions of Proposition 3.12.

LEMMA 4.29. *Let $(\bar{X},\bar{k})$ be the solution to* (3.5) *with initial condition $\bar{X}_0$ with law ν_0. Let $(X^{(N)}, k^{(N)})$ be a solution to the interacting particle system* (3.4) *with initial condition $X^{(N)} = (X_0^{1,N},\dots,X_0^{N,N})$. Assume that $(\bar{X}_0, X_0^{(N)})$ is independent on the noise $W^{(N)} = (W^1,\dots,W^N)$. For every $t\in[0,T]$, we have $\mathbb{P}$-a.s.,*

$$\mathcal{W}_{2,\mathbb{R}}\left(Law(\bar{X}_t), \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}\right)^2 \leq C\left(1+\mathbb{E}\left[\left(\int_0^T \mathrm{d}|\bar{k}|_s\right)^2\right]+\frac{1}{N}\sum_{i=1}^N\left(\int_0^T \mathrm{d}|k^{i,N}|_s\right)^2\right) \cdot \mathcal{W}_{2,C([0,T],\mathbb{R})}\left(Law(W), \frac{1}{N}\sum_{i=1}^N \delta_{W^i}\right)^2 + \mathcal{W}_{2,[0,1]}\left(\nu^0, \frac{1}{N}\sum_{i=1}^N \delta_{X_0^{i,N}}\right)^2.$$

Proof. For simplicity of notation, we take $\sigma = 1$ (the argument is the same for general $\sigma \in \mathbb{R}$). Call $\nu := \mathrm{Law}(W)$ the Wiener measure on $C_t = C([0,T],[0,1])$. For a fixed $\omega \in \Omega$, we consider the empirical measure $L^{N,\omega} := \frac{1}{N}\sum_{i=1}^N \delta_{(W^i(\omega), X_0^{i,N}(\omega))}$ as a law on $E = (C_t \times [0,1], \mathcal{B}(C_t)\times\mathcal{B}([0,1]))$. Let $P^\omega \in \mathcal{P}(E\times E)$ be any coupling of $\nu\otimes\nu_0$ and $L^{N,\omega}$. It is easy to verify that P^ω can be seen as a measure on $\Omega^\omega := (E \times \{(W^1(\omega), X_0^{1,N}(\omega)), \dots, (W^N(\omega), X_0^{N,N}(\omega))\})$, endowed with the product σ-algebra $\mathcal{A}^\omega := \mathcal{B}(C_t)\times 2^{\{(W^1(\omega),X_0^{1,N}(\omega)),\dots,(W^N(\omega),X_0^{N,N}(\omega))\}}$. Indeed, for every Borel bounded test function $\varphi : E\times E \to \mathbb{R}$,

$$\begin{aligned}
P^\omega(\varphi) &= \int_{E\times E} \varphi(x,y)P^\omega(\mathrm{d}x,y)L^{N,\omega}(\mathrm{d}y)\\
&= \frac{1}{N}\sum_{i=1}^N \int_E \varphi(x,(W^i(\omega),X_0^{i,N}(\omega))P^\omega\left(\mathrm{d}x, \mathrm{d}\left(W^i(\omega),X_0^{i,N}(\omega)\right)\right)\\
&= \int_{\Omega^\omega} \varphi(x,y)P^\omega(\mathrm{d}x,y)L^{N,\omega}(\mathrm{d}y).
\end{aligned}$$

On the space $(\Omega^\omega, \mathcal{A}^\omega, P^\omega)$ we define the projections $(\Pi^1,\Pi^1_0) \sim \nu\otimes\nu_0$ and $(\Pi^2,\Pi^2_0)\sim L^{N,\omega}$ on the first and the second marginal space, respectively (in particular, $\Pi^1\sim\nu$ and $\Pi^2 \sim \frac{1}{N}\sum_{i=1}^N \delta_{W^i(\omega)}$). Since the law of Π^1 is the Wiener measure ν, we have that Π^1 is a Brownian motion, and if we plug it as the driver of (4.18) we obtain a strong unique solution $(\bar{X},\bar{k})$ thanks to Theorem 3.9.

Let $(X^{(N)}, k^{(N)})$ be the solution of (3.4) given by Proposition 3.5. There exists a set of full measure $\Omega_0\subset\Omega$ such that for every $\omega\in\Omega_0$ and every $1\le i\le N$, $(X^{i,N}(\omega), k^{i,N}(\omega))$ satisfies (3.4). Defining $(\tilde{X},\tilde{k})(W^i(\omega),X_0^{i,N}(\omega)) := (X_t^{i,N}, k_t^{i,N})(\omega)$, we have that, for every $t\in[0,T]$, $\mathbb{E}_{P^\omega}[\tilde{k}_t] = \frac{1}{N}\sum_{j=1}^N k_t^{j,N}(\omega)$ and

$$\mathrm{d}\tilde{X}_t = (\mu(\tilde{X}_t) - \mathbb{E}_{P^\omega}[\mu(\tilde{X}_t)])\mathrm{d}t + \mathrm{d}q_t + \mathrm{d}\Pi^2_t - \mathrm{d}\tilde{k}_t - \mathrm{d}\mathbb{E}_{P^\omega}[\Pi^2_t] + \mathrm{d}\mathbb{E}_{P^\omega}[\tilde{k}_t] \quad \text{on } \Omega_0.$$

We define $b(\bar{X}_t) := \mu(\bar{X}_t) - \mathbb{E}_{P^\omega}[\mu(\bar{X}_t)]$ and we estimate the following:

$$\begin{aligned}
&\frac{1}{2}\mathrm{d}(\bar{X}_t - \tilde{X}_t - (\Pi^1_t-\Pi^2_t) + \mathbb{E}_{P^\omega}[\Pi^1_t-\Pi^2_t] + \int_0^t [b(\bar{X}_s)-b(\tilde{X}_s)]\mathrm{d}s)^2\\
&\quad= -(\bar{X}_t-\tilde{X}_t)\mathrm{d}\bar{k}_t + (\bar{X}_t-\tilde{X}_t)\mathrm{d}\tilde{k}_t\\
&\qquad+ (\bar{X}_t-\tilde{X}_t)\mathrm{d}\mathbb{E}_{P^\omega}[\bar{k}_t] - (\bar{X}_t-\tilde{X}_t)\mathrm{d}\mathbb{E}_{P^\omega}[\tilde{k}_t] + (\Pi^1_t-\Pi^2_t)\mathrm{d}\bar{k}_t - (\Pi^1_t-\Pi^2_t)\mathrm{d}\tilde{k}_t\\
&\qquad- (\Pi^1_t-\Pi^2_t)\mathrm{d}\mathbb{E}_{P^\omega}[\bar{k}_t] + (\Pi^1_t-\Pi^2_t)\mathrm{d}\mathbb{E}_{P^\omega}[\tilde{k}_t] - \mathbb{E}_{P^\omega}[\Pi^1_t-\Pi^2_t]\mathrm{d}\bar{k}_t + \mathbb{E}_{P^\omega}[\Pi^1_t-\Pi^2_t]\mathrm{d}\tilde{k}_t\\
&\qquad+ \mathbb{E}_{P^\omega}[\Pi^1_t-\Pi^2_t]\mathrm{d}\mathbb{E}_{P^\omega}[\bar{k}_t] - \mathbb{E}_{P^\omega}[\Pi^1_t-\Pi^2_t]\mathrm{d}\mathbb{E}_{P^\omega}[\tilde{k}_t]\\
&\qquad+ \left[\int_0^t [b(\bar{X}_s)-b(\tilde{X}_s)]\mathrm{d}s\right]\mathrm{d}(\mathbb{E}_{P^\omega}[\bar{k}_t] - \mathbb{E}_{P^\omega}[\tilde{k}_t]).
\end{aligned}$$

The first and second terms on the RHS are always negative by the conditions on the boundaries. If we take expectation under P^ω on both sides, we have that the third and fourth terms on the RHS vanish, because $\mathbb{E}_{P^\omega}[\bar{X}_t] = \mathbb{E}_{P^\omega}[\tilde{X}_t] = q_t$. Similarly, the expectation of the last term vanishes. Hence, we have that

$$\mathbb{E}_{P^\omega}\left[|\bar{X}_t - \tilde{X}_t - (\Pi^1_0 - \Pi^2_0) - (\Pi^1_t - \Pi^2_t) + \mathbb{E}_{P^\omega}[\Pi^1_t - \Pi^2_t] + \int_0^t [b(\bar{X}_t) - b(\tilde{X}_s)]\mathrm{d}s|^2 \right]$$
$$\leq 2\mathbb{E}_{P^\omega}\left[\sup_{t\in[0,T]} |\Pi^1_t - \Pi^2_t|^2\right] \left(\mathbb{E}_{P^\omega}\left[\left(\int_0^T \mathrm{d}|\bar{k}|_s\right)^2\right] + \mathbb{E}_{P^\omega}\left[\left(\int_0^T \mathrm{d}|\tilde{k}|_s\right)^2\right]\right)$$
$$+ \mathbb{E}_{P^\omega}[|\Pi^1_0 - \Pi^2_0|^2].$$

The proof is concluded by first using Gronwall's lemma and then choosing $P^\omega = P^\omega_W \otimes P^\omega_0$, where P^ω_W (resp., P^ω_0) is the optimal coupling in $\mathcal{W}_{2,C_t}(\nu, \frac{1}{N}\sum_{i=1}^N \delta_{W^i})$ (resp., $\mathcal{W}_{2,[0,1]}(\nu_0, \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_t})$). ◻

Thanks to the previous proposition, it is immediate to derive the convergence of the particle system to the McKean–Vlasov equation, provided that we have convergence at time 0 and a bound on the second moment of k^N.

Proof of Proposition 3.12. By Lemma 4.29 and using the Hölder inquality, we have

$$\mathbb{E}\left[\sup_{t\in[0,T]} \mathcal{W}_{2,[0,1]}\left(\mathrm{Law}(\bar{X}_t), \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_t}\right)\right]$$
$$\leq C\left(1 + \mathbb{E}\left[\left(\int_0^T \mathrm{d}|\bar{k}|_s\right)^2\right] + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\left(\int_0^T \mathrm{d}|k^{i,N}|_s\right)^2\right]\right)^{\frac{1}{2}}$$
$$\cdot\, \mathbb{E}\left[\mathcal{W}_{2,C([0,T],\mathbb{R})}\left(\mathrm{Law}(W), \frac{1}{N}\sum_{i=1}^N \delta_{W^i}\right)^2\right]^{\frac{1}{2}}$$
$$+ C\mathbb{E}\left[\mathcal{W}_{2,[0,1]}\left(\mathrm{Law}(\bar{X}_0), \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_0}\right)\right].$$

The first term on the RHS is uniformly bounded in N thanks to Lemma 4.7 and the 2-integrability of $\|\bar{k}\|_{BV}$ in Lemma 4.21. The empirical measure of independent random variables distributed as the Wiener measure converges in the Wasserstein metric to the Wiener measure as $O(1/\sqrt{\log(N)})$; see [BLG14]. The Wasserstein distance of the initial conditions converges faster. Remember $X^{i,N}_0 = Y^i + \sum_{j=1}^N \delta_{Y^j} + q(0)$, where $(Y^i)_{i\in\mathbb{N}}$ is a family of i.i.d. random variables. We see that the speed of convergnce of $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_0}$ is the same as the speed of convergence of $\sum_{i=1}^N \delta_{Y^j}$, which is $1/\sqrt{N}$; see [FG15]. For a fixed $\omega \in \Omega$, taking an optimal coupling $m = m(\omega)$ between $\mathrm{Law}(\bar{X}_0)$ and $\frac{1}{N}\sum_{i=1}^N \delta_{Y^i}$, we have that $(x, y - \mathbb{E}_m[y] - \mathbb{E}_m[x])_\# m(dx, dy)$ is a coupling between $\mathrm{Law}(\bar{X}_0)$ and $\frac{1}{N}\sum_{i=1}^N \delta_{X^i_0}$. We can compute

$$\mathcal{W}_{2,[0,1]}\left(\mathrm{Law}(\bar{X}_0), \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_0}\right)^2 \leq \mathbb{E}_m\left[|X - Y + \mathbb{E}_m Y - \mathbb{E}_m X|^2\right] = \mathrm{Var}_m(X - Y)$$
$$\leq \mathbb{E}_m|X - Y|^2 = \mathcal{W}_{2,[0,1]}\left(\mathrm{Law}(\bar{X}_0), \frac{1}{N}\sum_{i=1}^N \delta_{Y^i,}\right)^2.$$

Taking the square roots and the expectation under $\mathbb{P}$ concludes the proof. ◻

5. Appendix: Proof of Proposition 3.5. The system (3.4) can be seen as an SDE on the moving domain $H_t \cap [0,1]^N$, where

$$H_t = \left\{ x \in \mathbb{R}^N \mid \frac{1}{N}\sum_{i=1}^N x^i = q(t) \right\},$$

with normal boundary conditions. Indeed, formally, for each $i = 1, \dots N$ and $m = 0, 1$, on the boundary $x^i = m$, the direction of reflection $(-1)^m(e_i - N^{-1}(1, \dots 1))$ (e^i being the ith vector of the canonical basis) is orthogonal to the face $H_t \cap \{x \mid x^i = m\}$. Here we use this fact to show well-posedness of the system (3.4).

We introduce some notation. In the following, we fix N and omit the superscripts N and ϵ in the notation. We call $H = \{x \in \mathbb{R}^N \mid \frac{1}{N}\sum_{i=1}^N x^i = 0\}$, $1 = (1, 1, \dots 1) \in \mathbb{R}^N$, $\Pi : \mathbb{R}^N \to \mathbb{R}^N$ the projector on H, that is, $\Pi x = x - N^{-1}(x \cdot 1)1$. We take $A : H \to \mathbb{R}^{N-1}$ a linear isometry and we call $D_t = A\Pi(H_t \cap [0,1]^N)$. For $i = 1, \dots N$, $m = 0, 1$, we call $\partial_{i,m}[0,1]^N = \{x \in [0,1]^N \mid x^i = m\}$, $\partial_{i,m} D_t = A\Pi(H_t \cap \partial_{i,m}[0,1]^N)$, $\gamma_{i,m} = (-1)^m(e_i - N^{-1}1)$ the direction of reflection of (3.4) on the face $\partial_{i,m}[0,1]^N$, and $\nu_{i,m} = A\gamma_{i,m}$. For x in $\partial[0,1]^N = \cup_{i,m}\partial_{i,m}[0,1]^N$, we call

$$\Gamma(x) = \left\{ \sum_{i,m} c_{i,m}\gamma_{i,k} 1_{x\in\partial_{i,k}[0,1]^N} \mid c_{i,m} \geq 0 \ \forall i = 1, \dots N, m = 0, 1 \right\}.$$

Similarly, for y in $\partial D_t = \cup_{i,m}\partial_{i,m} D_t$, we call

$$N_t(y) = \left\{ \sum_{i,m} c_{i,m}\nu_{i,k} 1_{y\in\partial_{i,k} D_t} \mid c_{i,m} \geq 0 \ \forall i = 1, \dots N, m = 0, 1 \right\};$$

note that $N_t(y) = A\Pi(x)$ if x is in $\partial_{i,m}[0,1]^N$.

We consider the following SDE on D_t:

$$\begin{aligned} &\mathrm{d}Y_t = A\Pi b(t, A^{-1}Y_t + q(t)1)\mathrm{d}t + \Pi \mathrm{d}W_t + \mathrm{d}h_t, \\ &Y_t \in D_t \ \forall t, \quad P\text{-a.s.}, \\ &\mathrm{d}|h|_t = 1_{Y_t \in \partial D_t}\mathrm{d}|h|_t, \quad \mathrm{d}h_t = \nu_t \mathrm{d}|h|_t, \quad \nu_t \in N_t(Y_t), \end{aligned} \tag{5.1}$$

where (Y, h) is the solution, W is an N-dimensional Brownian motion with respect to a (complete, right-continuous) filtration $(\mathcal{F}_t)_t$, and b is the drift of the system (3.4). This is an SDE on a moving domain D_t with reflection at the boundary. As we will see, the SDE (5.1) is, up to the isometry A, the system (3.4).

LEMMA 5.1. *Under Condition* 3.2 *on q and the Lipschitz continuity of μ^ϵ, given a probability space $(\Omega, \mathcal{A}, P)$ and a Brownian motion W on a (complete right-continuous) filtration $(\mathcal{F}_t)_t$, there exists a unique (strong) solution to the SDE* (5.1).

Proof. The existence and uniqueness result is a consequence of [NO15, Theorem 1.7] for SDEs on moving domains with reflecting boundaries, provided that the assumptions of that theorem hold. We focus on two key assumptions, namely, (a) the fact that $N_t(y)$ is the cone of inward normal vectors of D_t at y, for every t and every $y \in \partial D_t$; (b) relation (1.16) in [NO15]. The other assumptions of [NO15, Theorem 1.7] are easy to verify.

Concerning assumption (a), we observe that, for each $i = 1, \dots N$, $m = 0, 1$, the vector $\gamma_{i,m}$ is the inward normal, in the $N-1$-dimensional convex polyhedron

$H_t \cap [0,1]^N$, of the corresponding face $H_t \cap \partial_{i,m}[0,1]^N$: indeed $\gamma_{i,m}$ belongs to H and, for every v in $H \cap \partial_{i,m}[0,1]^N$, we have $\gamma_{i,m} \cdot v = 0$. Since A is an isometry, the vector $\nu_{i,m} = A\gamma_{i,m}$ is the inward normal, in the convex polyhedron D_t, of the corresponding face $\partial_{i,m} D_t = A\Pi(H_t \cap \partial_{i,m}[0,1]^N)$. Now $N_t(y)$ is the convex cone generated by $\nu_{i,m}$, with i, m such that $y \in \partial_{i,m} D_t$. Hence $N_t(y)$ is the convex cone of inward normal vectors (in the sense of [NO15, Definition 2.2]); see, e.g., formula (4.23) in [Cos92].

Assumption (b) reads as follows. Define

$$a_{s,z}(\rho,\eta) = \max_{u\in\mathbb{R}^{N-1},|u|=1} \min_{s\le t\le t+\eta} \min_{y\in\partial D_t,|y-z|\le\rho} \min_{\nu\in N_t(y),|\nu|=1} (\nu \cdot u). \tag{5.2}$$

Condition (1.16) in [NO15] reads

$$\lim_{\eta\to 0} \lim_{\rho\to 0} \inf_{s\in[0,T]} \inf_{z\in\partial D_s} a_{s,z}(\rho,\eta) = a > 0. \tag{5.3}$$

In order to show this condition, we take

$$I = I_{s,z}(\rho,\eta) = \cup_{s\le t\le t+\eta} \cup_{y\in\partial D_t,|y-z|\le\rho} \{(i,m) \mid y \in \partial_{i,m} D_t\},$$
$$u_{s,z}(\rho,\eta) = c_I \sum_{(i,m)\in I_{s,z}(\rho,\eta)} \nu_{i,m},$$

where $c_I > 0$ is a positive constant such that $|u_{s,z}(\rho,\eta)| = 1$; note that $\min_I c_I = c_N > 0$. We also note that, for suitable $\rho_0 > 0$, $\eta_0 > 0$ (independent of s and z), for every $\rho < \rho_0$ and $\eta < \eta_0$, for every s and z, for every $j = 1,\dots N$, at most one element between $(j,0)$ and $(j,1)$ belongs to $I_{s,z}(\rho,\eta)$. Moreover, since the average of y is in $[\xi, 1-\xi]$ (for all $y \in \partial D_t$ for all t), y^i cannot be all 0, nor they can be all 1, hence $I_{s,z}$ cannot be $\{(1,0),(2,0),\dots(N,0)\}$ or $\{(1,1),(2,1),\dots(N,1)\}$. As a consequence,

$$\text{if } (i,m) \in I, \text{ then there exist at most } N-2 \text{ indices } j \ne i \text{ with } (j,m) \in I. \tag{5.4}$$

We compute the scalar products among $\nu_{i,m}$, using the isometry property of A:

$$|\nu_{i,m}|^2 = |\gamma_{i,m}|^2 = 1 - N^{-1},$$
$$\nu_{i,m} \cdot \nu_{j,m} = \gamma_{i,m} \cdot \gamma_{j,m} = -N^{-1} \quad \text{for } i \ne j,$$
$$\nu_{i,m} \cdot \nu_{j,n} = \gamma_{i,m} \cdot \gamma_{j,n} = N^{-1} \quad \text{for } i \ne j, m \ne n.$$

We call $\hat\nu_{i,m} = (1-N^{-1})^{-1/2}\nu_{i,m}$. For $\eta < \eta_0$ and $\rho < \rho_0$, we get by (5.4), for every (i,m) in $I = I_{s,z}(\rho,\eta)$,

$$\begin{aligned}\hat\nu_{i,m} \cdot u_I &= c_I(1-N^{-1})^{-1/2}\left(1 - N^{-1} - \sum_{(j,m)\in I, j\ne i} N^{-1} + \sum_{(j,n)\in I, j\ne i, m\ne n} N^{-1}\right)\\ &= c_I(1-N^{-1})^{-1/2}(1 - N^{-1} - (N-2)N^{-1})\\ &\ge c_N N^{-1}(1-N^{-1})^{-1/2}.\end{aligned}$$

Now, for every s, z, for every $t \in [s, s+\eta]$ and $y \in \partial D_t$ with $|y - z| \le \rho$, $N_t(y)$ is contained in the convex cone generated by $\hat\nu_{i,m}$, $(i,m) \in I_{s,z}(\eta,\rho)$. Therefore, for $\eta < \eta_0$ and $\rho < \rho_0$, for every s and z, we have

$$\nu \cdot u_I \ge c_N N^{-1}(1-N^{-1})^{-1/2} \quad \text{for every } \nu \text{ as in (5.2)},$$

and so $a_{s,z}(\rho,\eta) \ge c_N N^{-1}(1-N^{-1})^{-1/2} > 0$; in particular (5.3) holds. The proof is complete. ∎

Now we show that the SDE (5.1) is equivalent to the system (3.4). We introduce some notation. We take a Borel map

$$G : \{(t, y, v) \mid t \in [0, T], y \in \partial D_t, v \in N_t(y)\} \to [0, +\infty)^{N\times 2}, \quad (t, x, v) \mapsto (c_{i,m})_{i=1,\dots N, m=0,1},$$

such that $c_{i,m} = 0$ if y does not belong to $\partial_{i,m} D_t$, and

$$v = \sum_{(i,m), y\in\partial_{i,m} D_t} c_{i,m}\nu_{i,m}.$$

(Note that this map G exists but is not uniquely determined: indeed, if y belongs to $\partial_{i,0}D_t \cup \partial_{i,1}D_t$ for each i (that is, $y = A\Pi x$ for some x with $x^i \in \{0, 1\}$ for each i), then $\nu_{i,m}$ are not linearly independent.) For a solution (Y, h) to (5.1), with $\mathrm{d}h = \nu\mathrm{d}|h|$, we call

$$\begin{aligned} X^{Y,h}_t &= A^{-1}Y_t + q(t)1, \\ k^{Y,h}_t &= \int_0^t \sum_{(i,m)} G_{i,m}(r, Y_r, \nu_r) n(X^i_r) e_i 1_{X^i_r=m} \mathrm{d}|h|_r; \end{aligned}$$

recall that $n(m) = -(-1)^m$ is the outward normal of $[0, 1]$ in $m = 0, 1$.

LEMMA 5.2. *Assume that* (Y, h) *is a solution to* (5.1). *Then* $(X^{Y,h}, k^{Y,h})$ *is a solution to the system* (3.4).

Proof. Let (Y, h) be a solution to (5.1), and take $(X, k) = (X^{Y,h}, k^{Y,h})$. By the definition of $X^{Y,h}$ and $k^{Y,h}$, P-a.s., X has continuous paths with values in $[0, 1]$ and k has continuous paths, and, for each i, $|k^i|$ is concentrated on $\{t \mid X_i \in \{0, 1\}\}$ and has direction $n(X^i)$. Hence the second and third lines of (3.4) are satisfied. We have

$$\begin{aligned} \mathrm{d}h_t &= \sum_{(i,m)} G_{i,m}(t, Y_t, \nu_t) 1_{Y_t\in\partial_{i,m}D_t} \nu_{i,m} \mathrm{d}|h|_t \\ &= \sum_{(i,m)} G_{i,m}(t, Y_t, \nu_t) 1_{X^i_t=m} A(-1)^m (e_i - N^{-1}1) \mathrm{d}|h|_t \\ &= -\sum_{(i,m)} G_{i,m}(t, Y_t, \nu_t) 1_{X^i_t=m} n(X^i_t) A e_i \mathrm{d}|h|_t + N^{-1} \sum_{(i,m)} G_{i,m}(t, Y_t, \nu_t) 1_{X^i_t=m} n(X^i_t) A1 \mathrm{d}|h|_t \\ &= -A\mathrm{d}k_t + N^{-1}A(\mathrm{d}k \cdot 1)1 = A\left(-\mathrm{d}k_t + N^{-1}\sum_i \mathrm{d}k^i_t 1\right). \end{aligned}$$

Hence, applying the transformation $X_t = A^{-1}Y_t + q(t)1$ to the first line of (5.1), we obtain the first line of (3.4). Therefore $(X^{Y,h}, k^{Y,h})$ satisfies (3.4). The proof is complete. ∎

By Lemmas 5.1 and 5.2, we get existence of a solution to (3.4).

Remark 5.3. We expect also the converse of Lemma 5.2 to hold, namely, if (X, k) solves (3.4), then $(Y = A\Pi X, h = -A\Pi k)$ solves (5.1). In particular, from this converse we would get uniqueness for (3.4). However, showing the third line of (5.1) is not immediate, hence we do not follow this strategy.

We conclude the proof of Proposition 3.5 by showing uniqueness for (3.4).

LEMMA 5.4. *Strong uniqueness (in* X*) holds for the SDE* (3.4).

Proof. The proof follows the line of Proposition 4.2, replacing the expectation with the empirical average. Let (X, k^X), (Y, k^Y) be two solutions to (3.4) with the same initial condition $X_0 = Y_0$. In this proof we call $\mathrm{d}K^X = \frac{1}{N}\sum_{j=1}^N[\mu(X^j)\mathrm{d}t - \mathrm{d}W^j + \mathrm{d}k^{X,j}]$ and similarly for K^Y. By the Itô formula for continuous semimartingales, we have, for every $i = 1, \dots N$,

$$\begin{aligned}
&\mathrm{d}|X^i - Y^i|^2 \\
&\quad = 2(X^i - Y^i)(-\mu(X^i) + \mu(Y^i))\mathrm{d}t + 2(X^i - Y^i)\mathrm{d}K^X - 2(X^i - Y^i)\mathrm{d}K^Y \\
&\qquad - 2(X^i - Y^i)\mathrm{d}k^{X,i} + 2(X^i - Y^i)\mathrm{d}k^{Y,i}.
\end{aligned}$$

The one-side Lipschitz condition of μ implies

$$(X^i - Y^i)(-\mu(X^i) + \mu(Y^i)) \le c|X^i - Y^i|^2,$$

and the orientation of k (as the outward normal) implies

$$-\int_0^t (X^i - Y^i)\mathrm{d}k^{X,i} \le 0$$

and similarly for $(X^i - Y^i)\mathrm{d}k^{Y,i}$. For the addends with K, we average over i and use that K does not depend on i and that $\frac{1}{N}\sum_i X_t^i = \frac{1}{N}\sum_i Y_t^i = q(t)$: we obtain

$$\frac{1}{N}\sum_i (X^i - Y^i)\mathrm{d}K^X = 0$$

and similarly for $(X^i - Y^i)\mathrm{d}K^Y$. Putting it all together, we get

$$\frac{1}{N}\sum_i |X_t^i - Y_t^i|^2 \le C\int_0^t \frac{1}{N}\sum_i |X_r^i - Y_r^i|^2 \mathrm{d}r.$$

We conclude by the Gronwall inequality that $\frac{1}{N}\sum_i |X_t^i - Y_t^i|^2 = 0$, that is, $X = Y$. The proof is complete. ∎

Acknowledgments. We thank Boualem Djehiche, Jean-Francois Jabir, John Schoenmakers, and Andreas Sojmark for pointing out relevant references on McKean–Vlasov SDEs on bounded domains and suggesting possible alternative approaches to our problem.

REFERENCES

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, Springer, New York, 2008.

[Aid16] S. Aida, *Rough Differential Equations Containing Path-Dependent Bounded Variation Terms*, arXiv:1608.03083, 2016.

[Bar20] C. L. Barnes, *Hydrodynamic limit and propagation of chaos for Brownian particles reflecting from a Newtonian barrier*, Ann. Appl. Probab., 30 (2020), pp. 1582–1613.

[BCCdRH20] P. Briand, P. Cardaliaguet, P.-É. Chaudru de Raynal, and Y. Hu, *Forward and backward stochastic differential equations with normal constraints in law*, Stochastic Process. Appl., 130 (2020), pp. 7021–7097.

[BCD20] I. Bailleul, R. Catellier, and F. Delarue, *Solving mean field rough differential equations*, Electron. J. Probab., 25 (2020), 21.

[BCdRGL20] P. Briand, P.-É. Chaudru de Raynal, A. Guillin, and C. Labart, *Particles systems and numerical schemes for mean reflected stochastic differential equations*, Ann. Appl. Probab., 30 (2020), pp. 1884–1909.

[BEH18] P. Briand, R. Elie, and Y. Hu, *BSDEs with mean reflection*, Ann. Appl. Probab., 28 (2018), pp. 482–510.

[BJ11] M. Bossy and J.-F. Jabir, *On confined McKean Langevin processes satisfying the mean no-permeability boundary condition*, Stochastic Process. Appl., 121 (2011), pp. 2751–2775.

[BJ15] M. Bossy and J.-F. Jabir, *Lagrangian stochastic models with specular boundary condition*, J. Funct. Anal., 268 (2015), pp. 1309–1381.

[BJ18] M. Bossy and J.-F. Jabir, *Particle approximation for Lagrangian stochastic models with specular boundary condition*, Electron. Commun. Probab., 23 (2018), 15.

[BLG14] E. Boissard and T. Le Gouic, *On the mean speed of convergence of empirical and occupation measures in wasserstein distance*, Ann. Inst. Henri Poincaŕe Probab. Stat., 50, (2014), pp. 539–563.

[CCP11] M. J. Cáceres, J. A. Carrillo, and B. Perthame, *Analysis of nonlinear noisy integrate & fire neuron models: Blow-up and steady states*, J. Math. Neurosci., 1 (2011), 7.

[CD18] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications.* I, Mean Field FBSDEs, Control, and Games, Probab. Theory Stoch. Model. 83, Springer, Cham, 2018.

[CDFM20] M. Coghi, J.-D. Deuschel, P. K. Friz, and M. Maurelli, *Pathwise McKean-Vlasov theory with additive noise*, Ann. Appl. Probab., 30 (2020), pp. 2355–2392.

[CL15] T. Cass and T. Lyons, *Evolving communities with individual preferences*, Proc. Lond. Math. Soc. (3), 110 (2015), pp. 83–107.

[Cos92] C. Costantini, *The Skorohod oblique reflection problem in domains with corners and application to stochastic differential equations*, Probab. Theory Related Fields, 91 (1992), pp. 43–70.

[DEH19] B. Djehiche, R. Elie, and S. Hamadène, *Mean-Field Reflected Backward Stochastic Differential Equations*, arXiv:1911.06079, 2019.

[DGH11] W. Dreyer, C. Guhlke, and M. Herrmann, *Hysteresis and phase transition in many-particle storage systems*, Contin. Mech. Thermodyn., 23 (2011), pp. 211–231.

[DGHT19] A. Deya, M. Gubinelli, M. Hofmanová, and S. Tindel, *One-dimensional reflected rough differential equations*, Stochastic Process. Appl., 129 (2019), pp. 3261–3281.

[DHM+15] W. Dreyer, R. Huth, A. Mielke, J. Rehberg, and M. Winkler, *Global existence for a nonlocal and nonlinear Fokker-Planck equation*, Z. Angew. Math. Phys., 66 (2015), pp. 293–315.

[DI91] P. Dupuis and H. Ishii, *On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications*, Stochastics Stochastics Rep., 35 (1991), pp. 31–62.

[DIRT15] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré, *Global solvability of a networked integrate-and-fire model of McKean-Vlasov type*, Ann. Appl. Probab., 25 (2015), pp. 2096–2133.

[FG15] N. Fournier and A. Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Related Fields, 162 (2015), pp. 707–738.

[FR13] M. Ferrante and C. Rovira, *Stochastic differential equations with non-negativity constraints driven by fractional Brownian motion*, J. Evol. Equ., 13 (2013), pp. 617–632.

[GGM+18] C. Guhlke, P. Gajewski, M. Maurelli, P. K. Friz, and W. Dreyer, *Stochastic many-particle model for LFP electrodes*, Contin. Mech. Thermodyn., 30 (2018), pp. 593–628.

[HLSj19] B. Hambly, S. Ledger, and A. Søjmark, *A McKean-Vlasov equation with positive feedback and blow-ups*, Ann. Appl. Probab., 29 (2019), pp. 2338–2373.

[Jab17] J.-F. Jabir, *Diffusion Processes with Weak Constraint Through Penalization Approximation*, arXiv:1704.01505, 2017.

[Kol07] V. N. Kolokoltsov, *Nonlinear Markov semigroups and interacting Lévy type processes*, J. Stat. Phys., 126 (2007), pp. 585–642.

[Li14] J. Li, *Reflected mean-field backward stochastic differential equations. Approximation and associated nonlinear PDEs*, J. Math. Anal. Appl., 413 (2014), pp. 47–68.

[Lio08] P.-L. Lions, *Lectures at collège de france*, 2008.

[LS84] P.-L. Lions and A.-S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*, Comm. Pure Appl. Math., 37 (1984), pp. 511–537.

[Ḿ96] S. Méléard, *Asymptotic behavior of some interacting particle systems; McKean-Vlasov and Boltzmann models*, in Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995), Lecture Notes in Math. 1627, Springer, Berlin, 1996, pp. 42–95.

[Men83] J.-L. Menaldi, *Stochastic variational inequality for reflected diffusion*, Indiana Univ. Math. J., 32 (1983), pp. 733–744.

[NO15] K. Nyström and T. Önskog, *Remarks on the Skorohod problem and reflected Lévy driven SDEs in time-dependent domains*, Stochastics, 87 (2015), pp. 747–765.

[RY99] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Grundlehren Math. Wiss. 293, Springer-Verlag, Berlin, 1999.

[Son07] H. M. Soner, *Stochastic representations for nonlinear parabolic PDEs*, in Handbook of Differential Equations: Evolutionary Equations. Vol. 3, Elsevier/North-Holland, Amsterdam, 2007, pp. 477–526.

[Szn84] A.-S. Sznitman, *Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated*, J. Funct. Anal., 56 (1984), pp. 311–336.

[Szn91] A.-S. Sznitman, *Topics in propagation of chaos*, in Ecole d'été de probabilités de Saint-Flour XIX-1989, Springer, Berlin, 1991, pp. 165–251.

[Tan79] H. Tanaka, *Stochastic differential equations with reflecting boundary condition in convex regions*, Hiroshima Math. J., 9 (1979), pp. 163–177.

[Tan84] H. Tanaka, *Limit theorems for certain diffusion processes with interaction*, in Stochastic Analysis (Katata/Kyoto, 1982), North-Holland Math. Library 32, North-Holland, Amsterdam, 1984, pp. 469–488.